

Analysis of Protein Structure and Intermolecular Interactions

Christopher Hogue

Samuel Lunenfeld Research Institute

Toronto.

hogue@mshri.on.ca

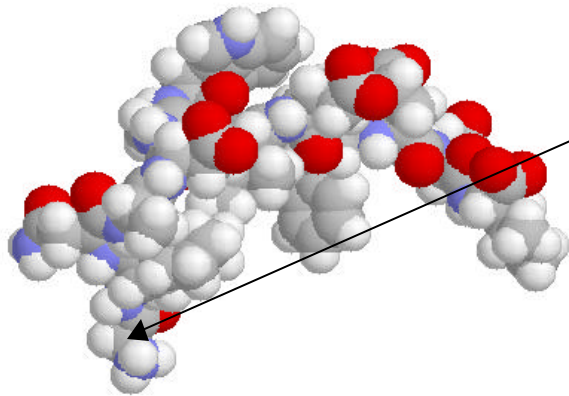
Bioinformatics

Testing biological hypotheses using
information technology and
computer science

A Sequence is an Abstraction

- Sequence is an implicit notation for the chemical graph of a standard biopolymer
 - Protein, DNA, RNA
 - Limited “standard” alphabet
- A protein is a heteropolymer of amino acids attached through condensation...

Protein Sequence - Database Abstraction



Biopolymer Sequence

GDKNADGWIEFEEL

Database of Sequences

Biopolymer molecule = string

Analyses - Similarity

Regular Expressions

BLAST

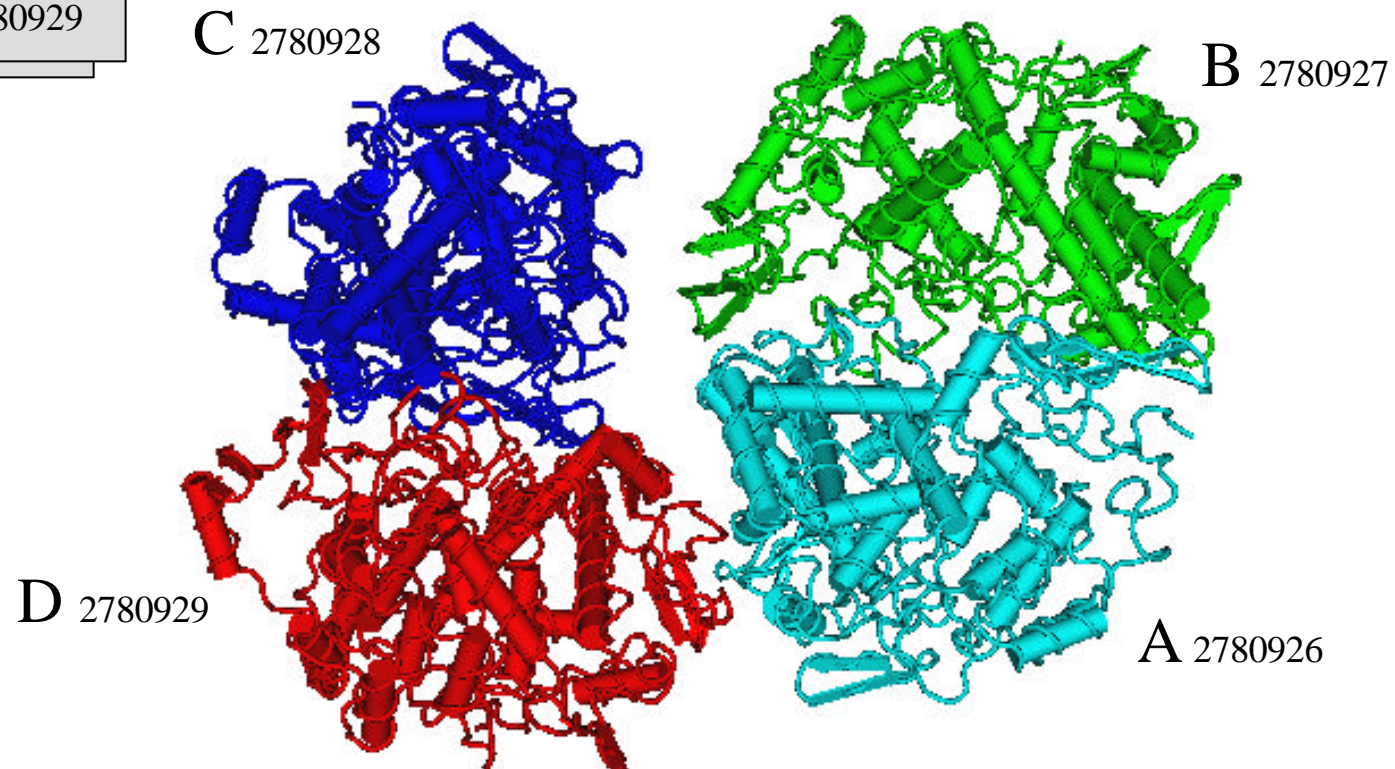
Evolution

Origin of Life

Link database


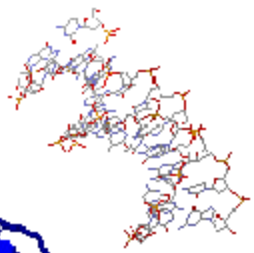

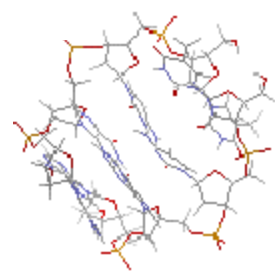
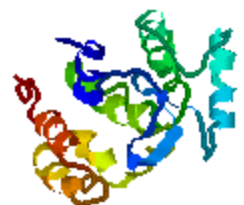
Structure	GI
4COX A	2780926
4COX B	2780927
4COX C	2780928
4COX D	2780929

What is in a protein structure?

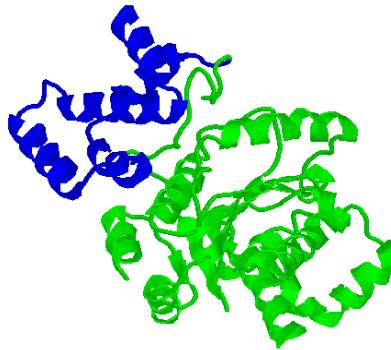


Cyclooxygenase-2 (Prostaglandin Synthase-2)

MMDB Database

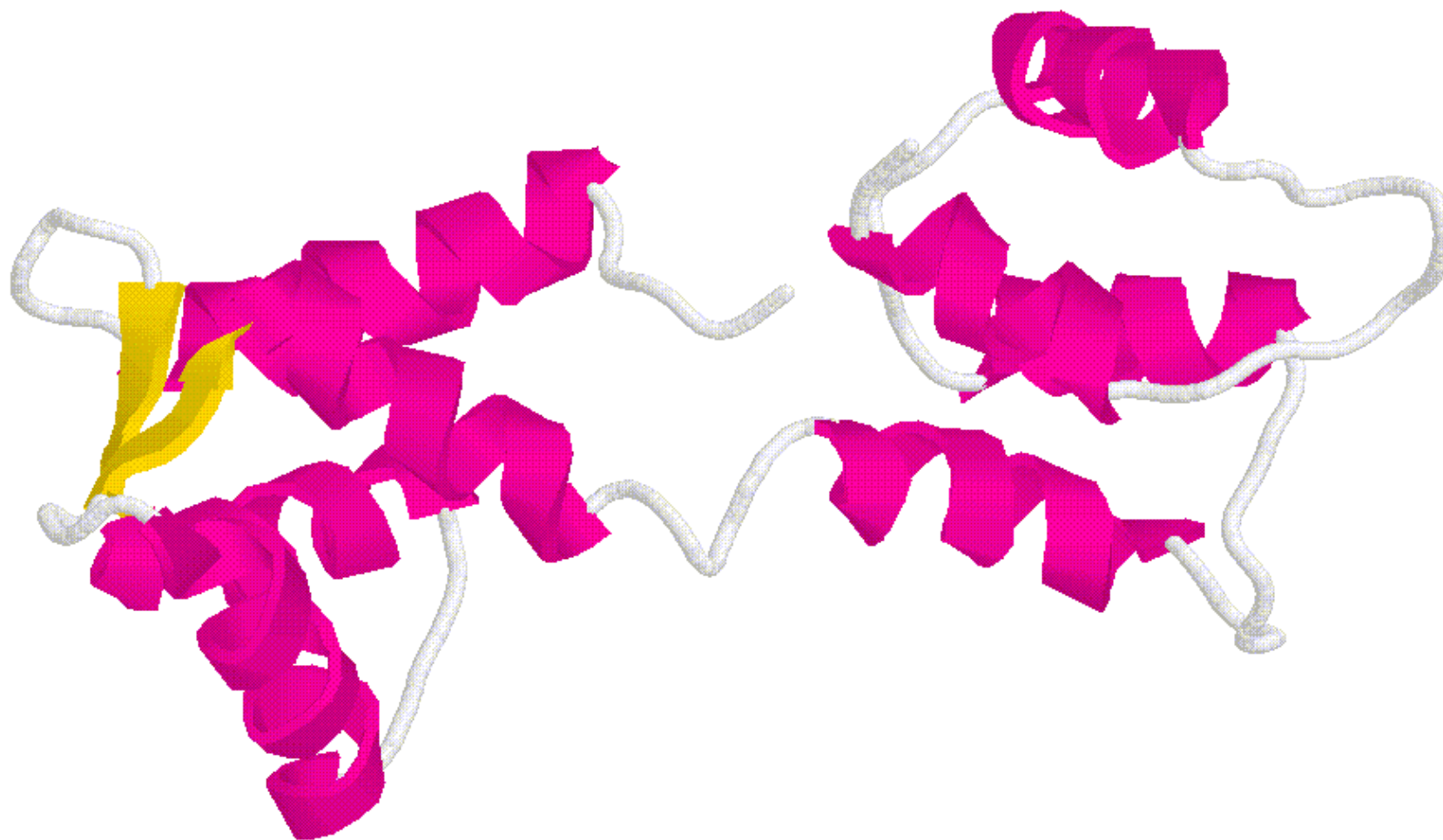
mmdbid	pdbid	models	molecules	structure
3543	177L	4	10	
130	178D	3	76	
3542	178L	4	11	
131	179D	8	2	
3541	179L	4	7	

two domains resolved: 1-245(green), 246-300 (blue)
unresolved: 300-400 (red)

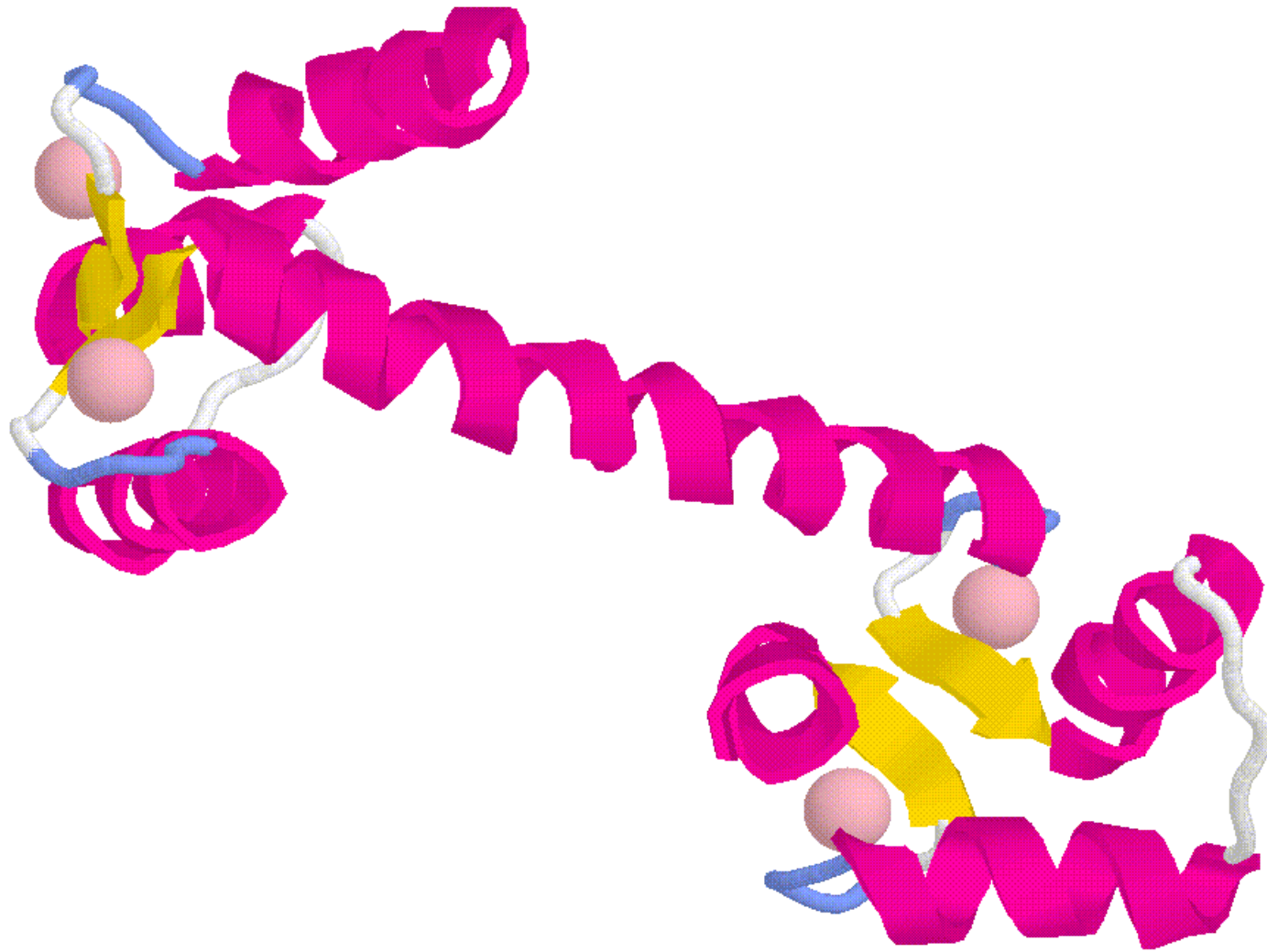
[illegible][illegible]

Protein Structure

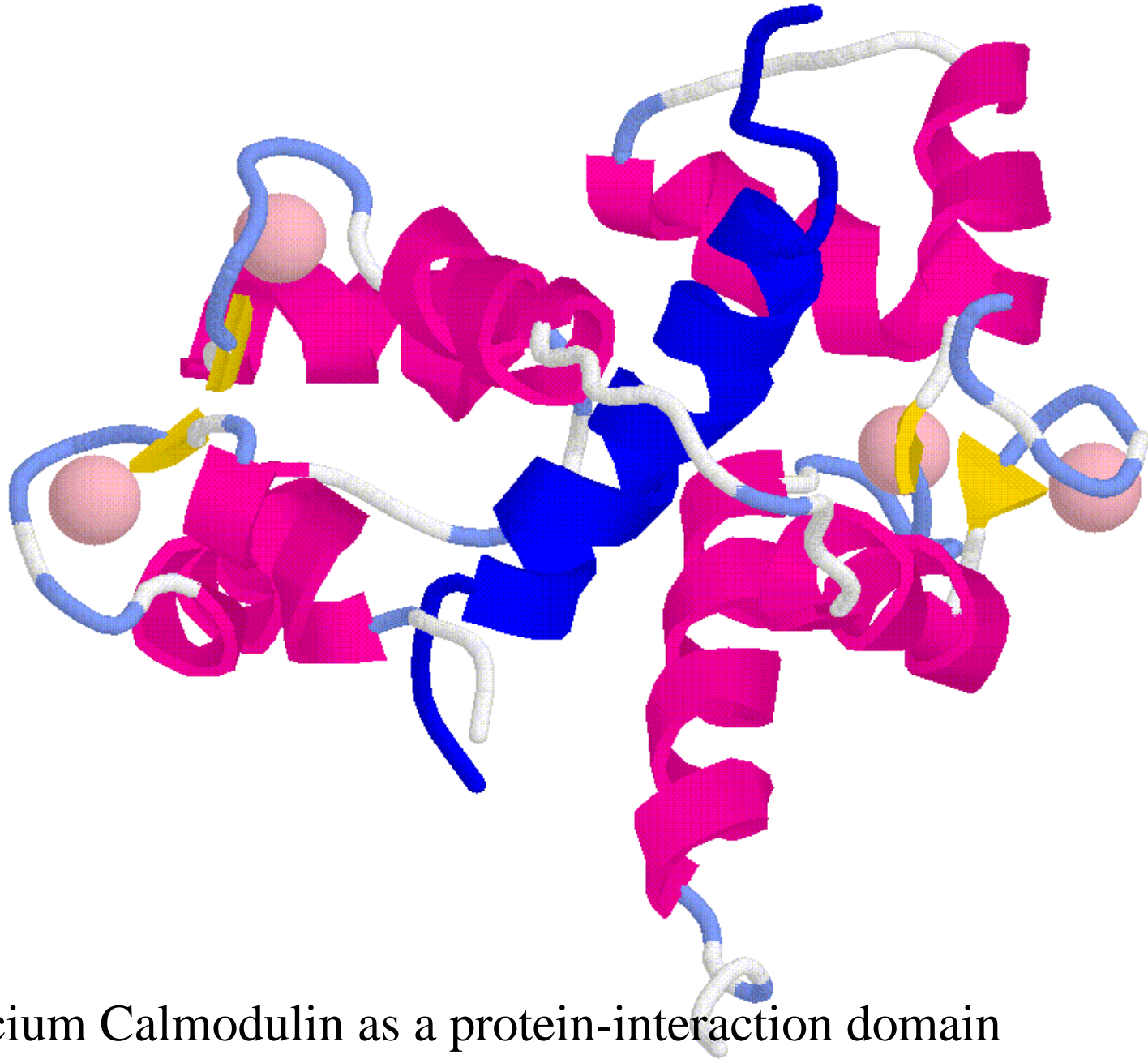
- Most proteins fold into one conformation but there is a dynamic range of conformational space available to protein structures.
- “Dynamics”



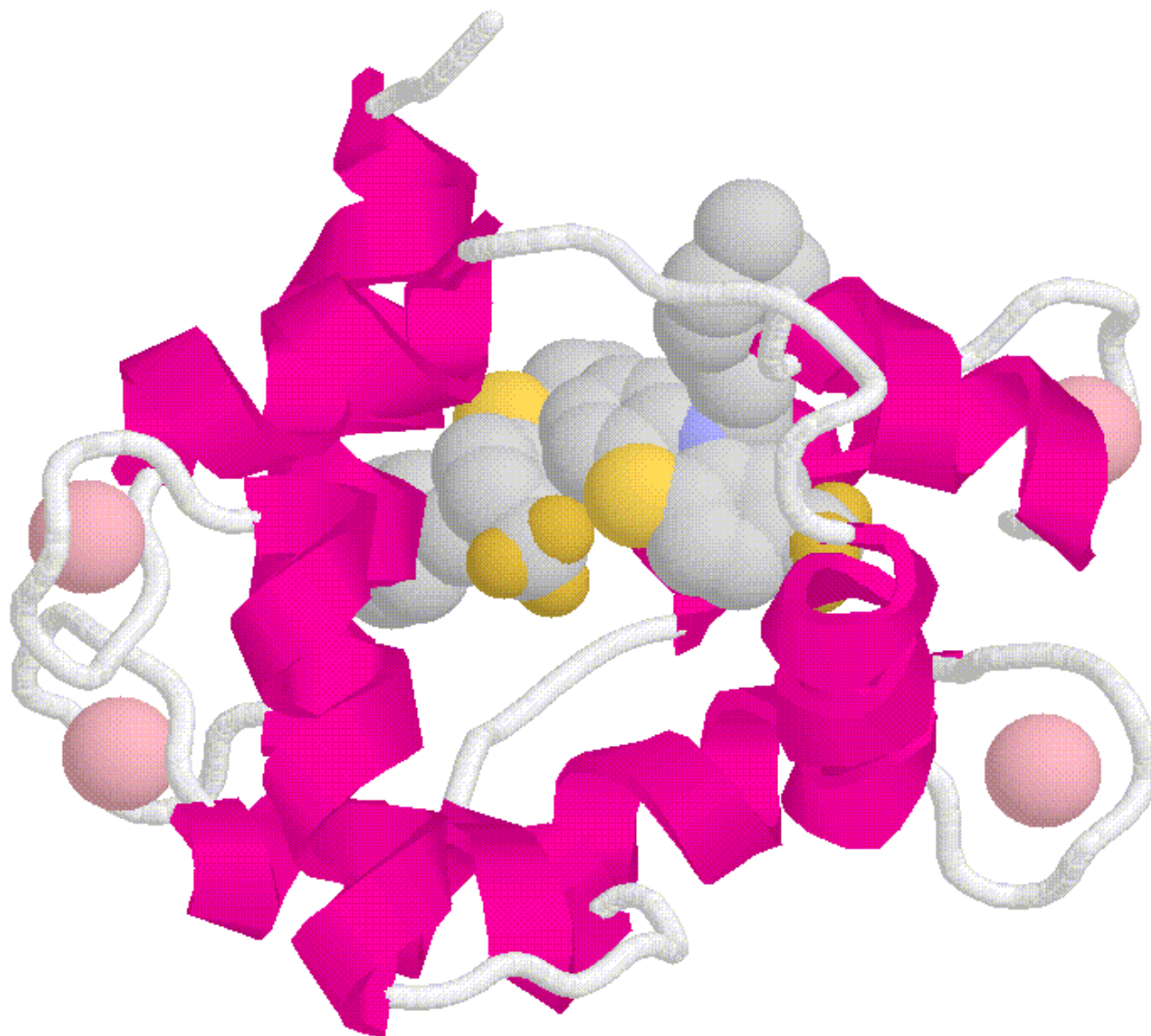
Apo (calcium free) calmodulin



Calcium bound calmodulin



Calcium Calmodulin as a protein-interaction domain



Calcium Calmodulin complexed with the inhibitor trifluoperazine

Amino Acid Sidechain Rotamers...

Leu



Lys



Met



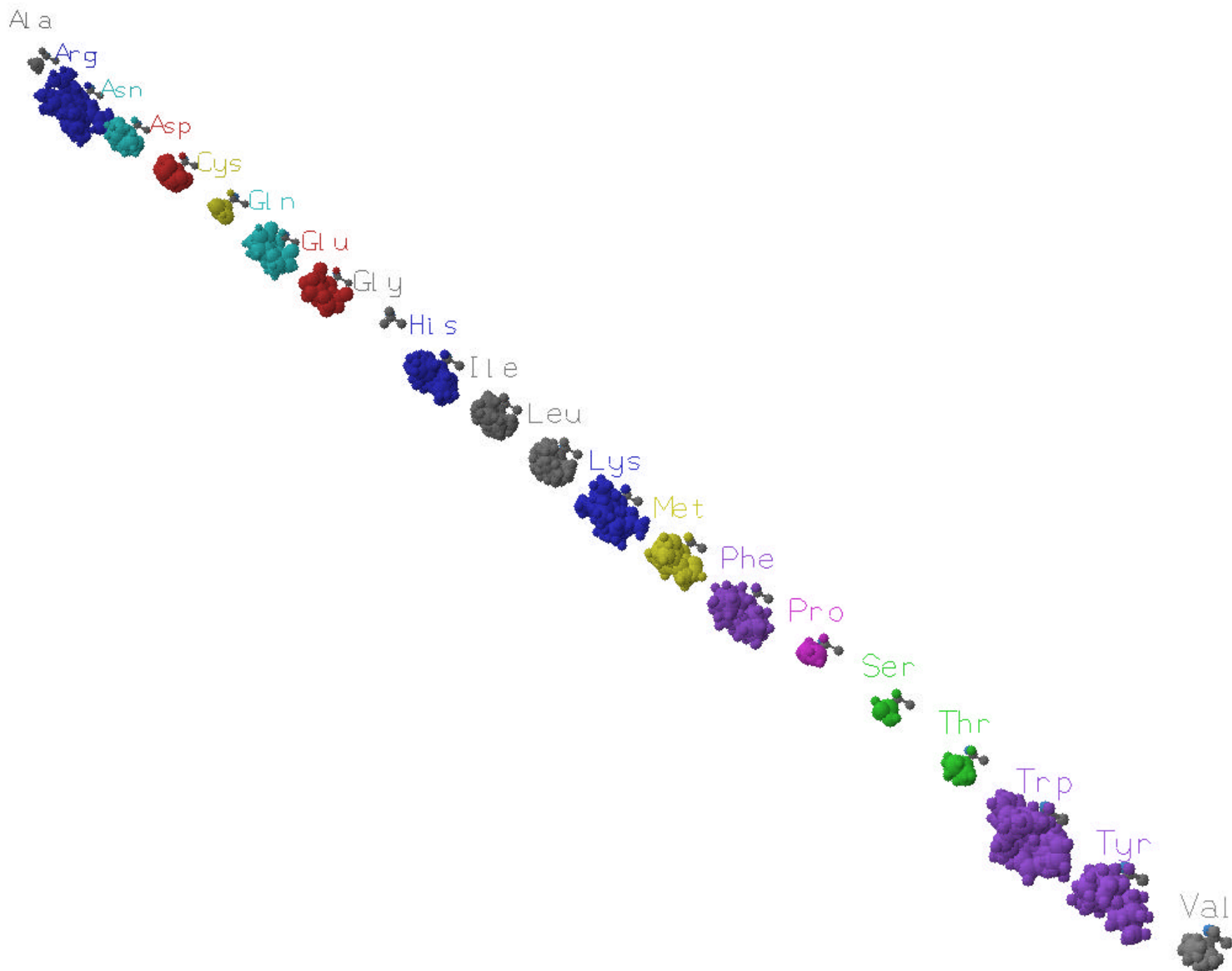
Phe



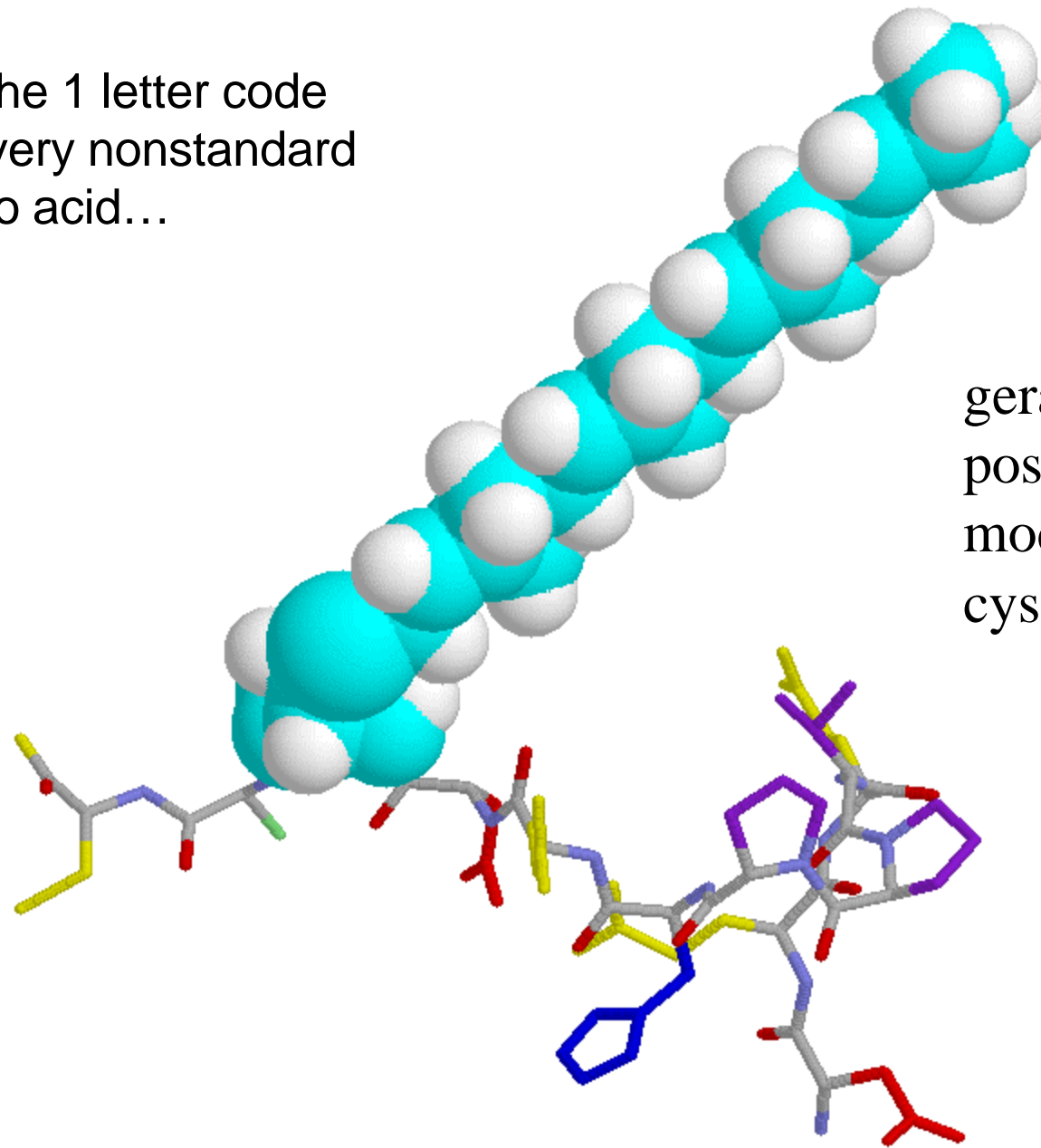
Pro



...



X is the 1 letter code
for every nonstandard
amino acid...



geranylgeranyl
post-translationally
modified
cysteine

Some Nonstandard Amino Acids...

METHYLATED

N-methyl-L-alanine	[A:meth_n]	Amino
N,N,N-trimethyl-L-alanine	[A:meth_n3]	Amino
omega-N,omega-N-dimethyl-L-arginine	[R:meth_n7]	Any
L-beta-methylthioaspartic acid	[D:meth_b]	Any
N5-methyl-L-glutamine	[Q:meth_n5]	Any
L-glutamic acid 5-methyl ester	[E:meth_o5]	Any
3'-methyl-L-histidine	[H:meth_n4]	Any
N6-methyl-L-lysine	[K:meth_1]	Any
N6,N6-dimethyl-L-lysine	[K:meth_2]	Any
N6,N6,N6-trimethyl-L-lysine	[K:meth_3]	Any
N-methyl-L-methionine	[M:meth]	Amino
N-methyl-L-phenylalanine	[F:meth]	Amino






PHOSPHORYLATED

omega-N-phospho-L-arginine	[R:po]	Any
L-aspartic 4-phosphoric anhydride	[D:po]	Any
S-phospho-L-cysteine	[C:po]	Any
1'-phospho-L-histidine	[H:po_e]	Any
3'-phospho-L-histidine	[H:po_d]	Any
O-phospho-L-serine	[S:po]	Any
O-phospho-L-threonine	[T:po]	Any
O4'-phospho-L-tyrosine	[Y:po]	Any

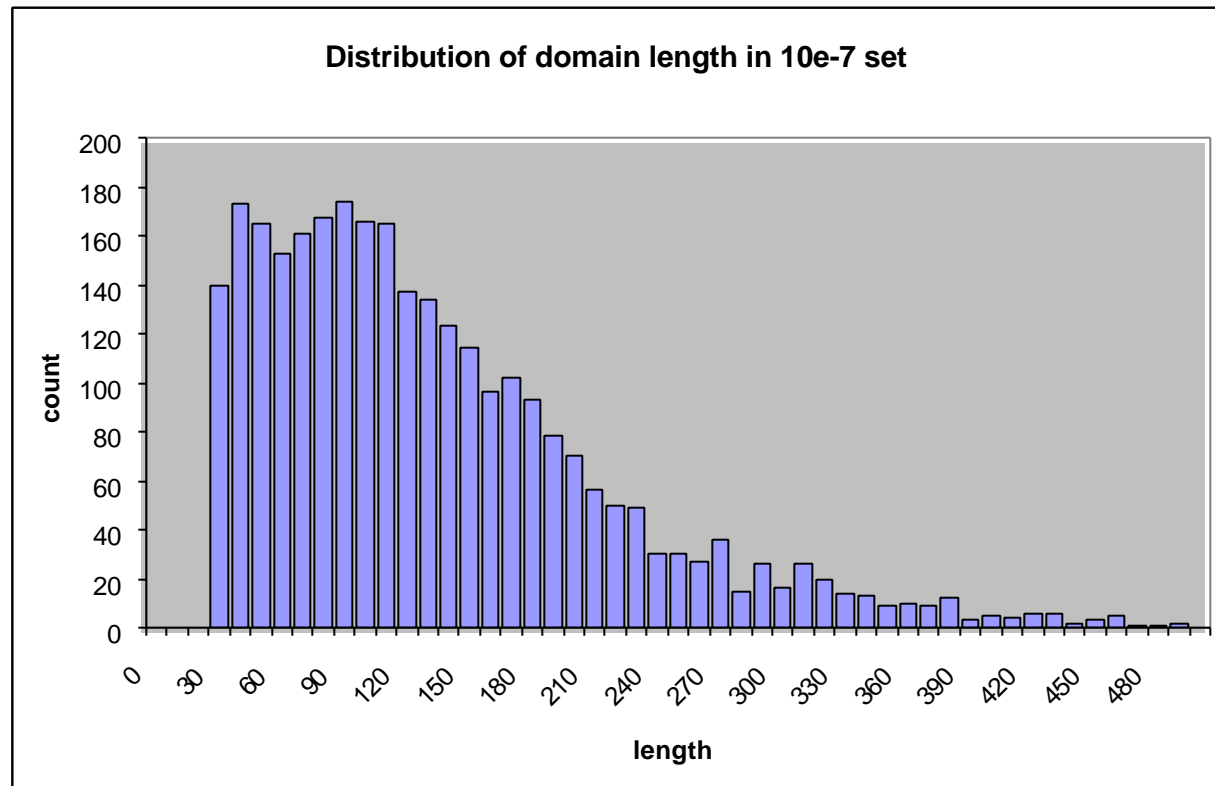
OTHER

L-selenocysteine	[C:sel]	Any
L-selenomethionine	[M:sel]	Any
L-3-oxoalanine	[S:oxal]	Any
2-pyrrolidone-5-carboxylic acid	[E:pyro]	Amino
L-glutamyl 5-glycerylphosphorylethanolamine	[E:gpe]	Any
2'-[3-carboxamido-3-(trimethylammonio)propyl]-L-histidine (diphthamide)	[H:diph]	Any
N6-biotinyl-L-lysine	[K:biotin]	Any
N6-(4-amino-2-hydroxybutyl)-L-lysine (hypusine)	[K:hypu]	Any
N6-retinal-L-lysine	[K:retin]	Any

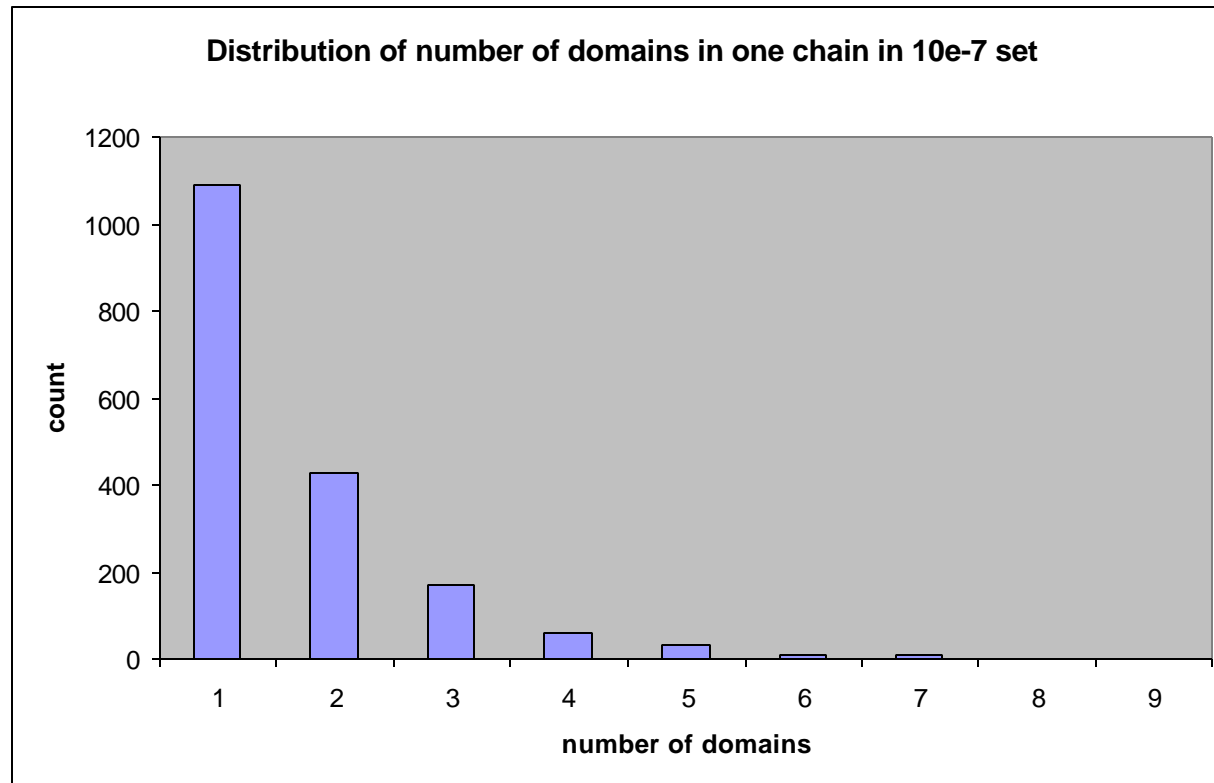
A Domain Definition

-  **A compact substructure of a protein based on the 3-D fold**
 -  **Defined without regard to the sequence conservation shared with other members of protein families**
 -  **Without regard to sequence continuity (i.e. topology)**
 -  **The domain division in MMDB is generated by the Vector Alignment Search Tool (VAST)**
 -  **VAST generates a domain division based on the compactness of adjacent secondary structure elements in space**
-

Domain Sizes

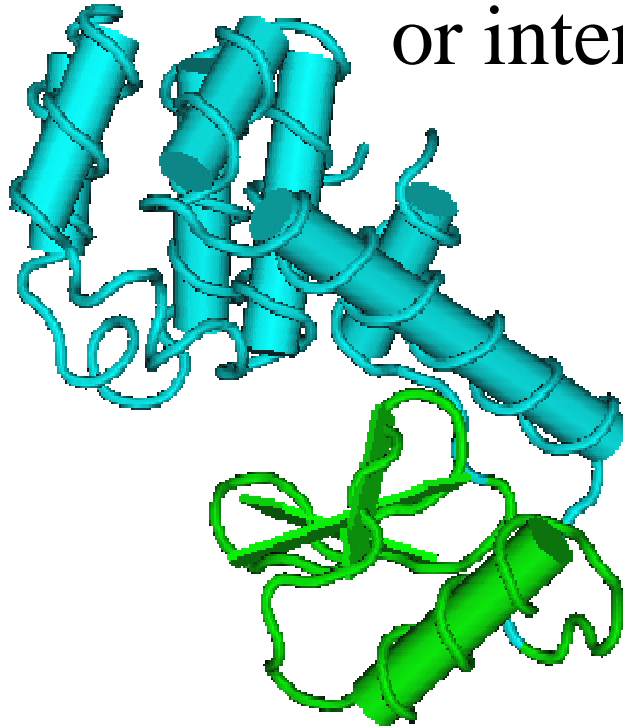


Domains in Protein Structures

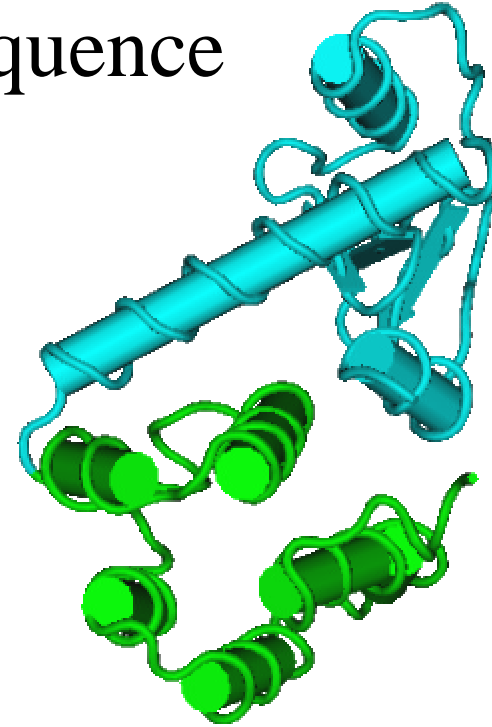


Very different from the genome perspective,
most human proteins are multidomain...

Domains are comprised of contiguous
or interrupted sequence



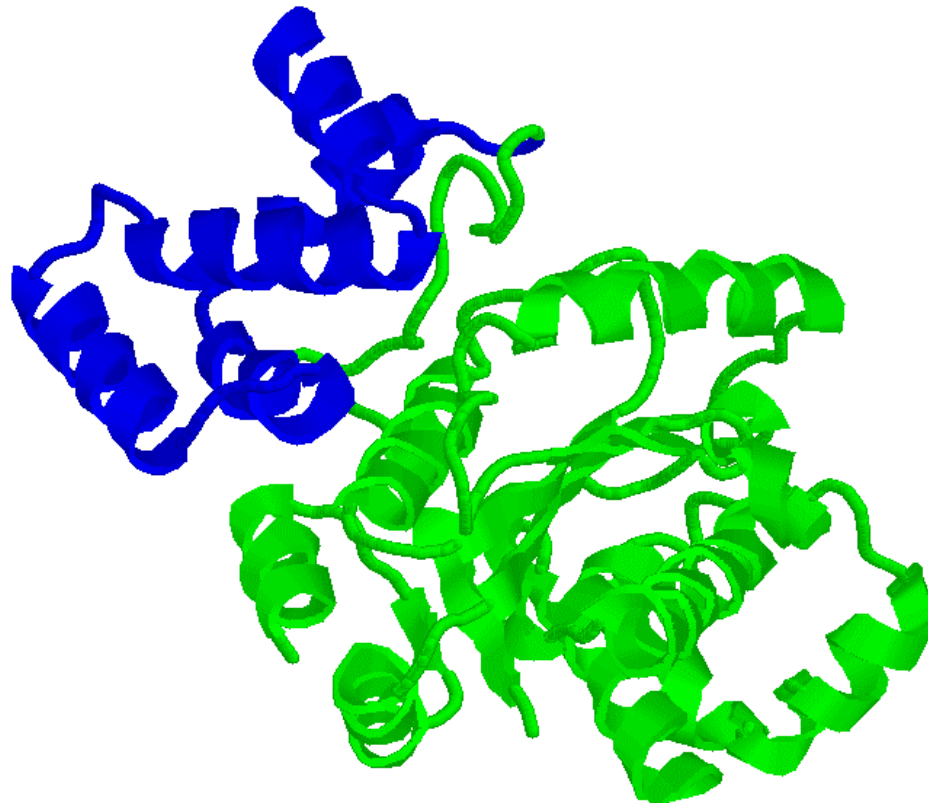
OR



	interrupted	contiguous
domains	6,493 (17%)	31,057 (83%)
chains	4,198 (18%)	18,975 (82%)
structures	2,711 (23%)	9,196 (77%)

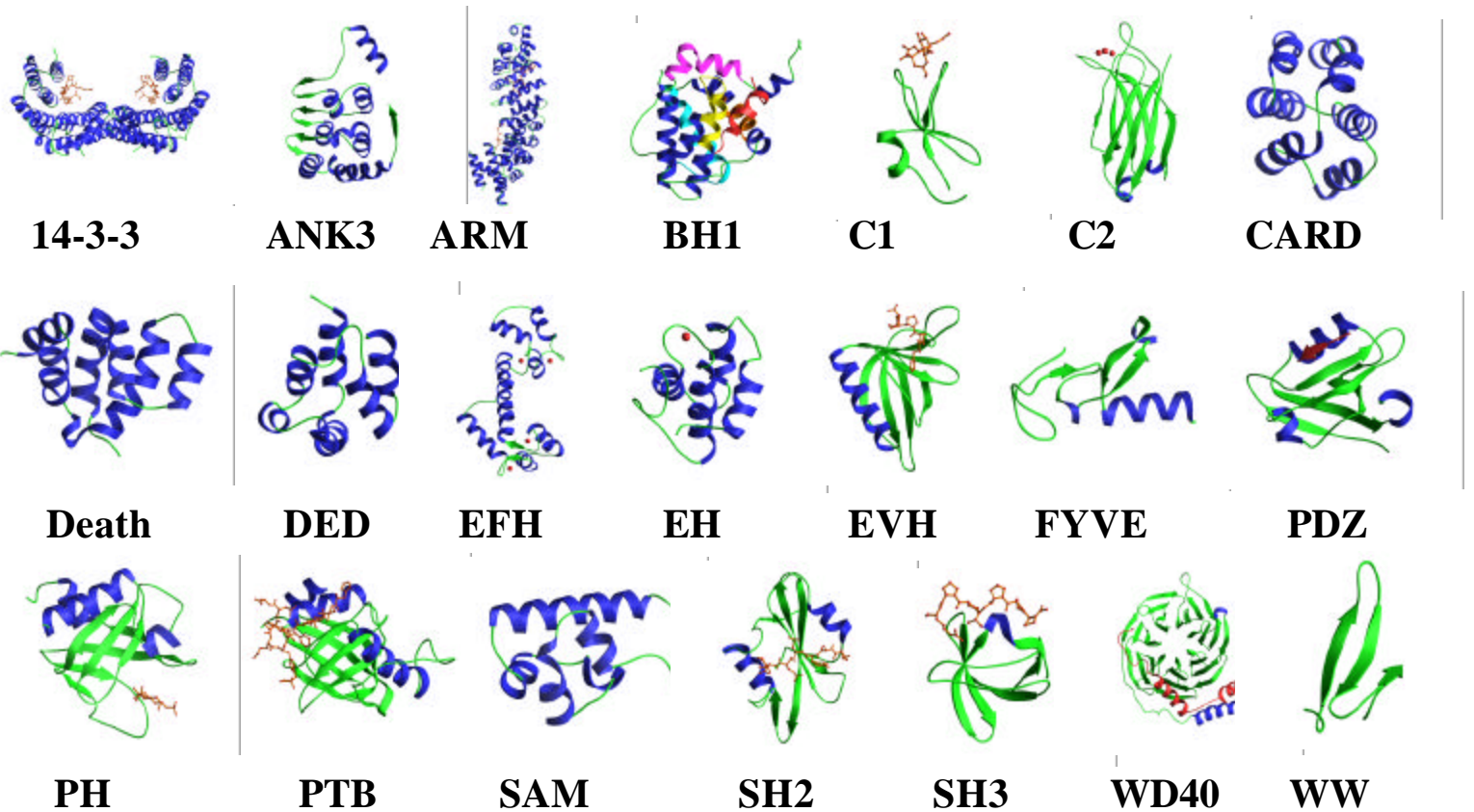
Tyrosyl tRNA synthetase

Proteins are modular and often functionally divisible.
Structural domains often correlate with protein functionality.



Rossman fold, a nucleotide binding domain, in green
tRNA recognition and binding domain in blue


Some Modular Protein Domains



Detecting Domains

- Most protein structure domains are known
- Many can be identified with PSSMs
- New tools - CDD (SMART, PFAM)

Most Human Proteins are Multidomain



The screenshot shows a Netscape browser window with the address bar displaying the URL: `ih.gov:80/entrez/viewer.cgi?save=0&cmd=&view=fasta&txt=on&val=2944396`. The main content area displays a protein sequence for the gene `g1|2944396|gb|AAC12931.1|`, identified as `guanine nucleotide exchange factor UNC-73A [Caenorhabditis elegans]`. The sequence is shown in a single-line FASTA format, starting with `MGDYDGKMTMKAEDILHVLRDGIAVLPGGRCRAGQAVIVCPSPREQPVNQDNLNRNVFLYLFEVTSKMAREK` and ending with `SLRNFARNSSYSFRVFAYNQVRVSAPSQCICICFDGTEM`. The browser's status bar at the bottom indicates "Document: Done".

```
>g1|2944396|gb|AAC12931.1| guanine nucleotide exchange factor UNC-73A [Caenorhabditis elegans]
MGDYDGKMTMKAEDILHVLRDGIAVLPGGRCRAGQAVIVCPSPREQPVNQDNLNRNVFLYLFEVTSKMAREK
GFLVVIDMRGKQTWINVRHILKALSSIESSTVQVFIIKPEKFWKQKAQMSLGTWDFEVEMISFESLIK
IIDSSHLPKTVGGSYPYDHEWLELRDLLEKWIWNITEIMEKLESVRREICEGEQPVVDVTTANAALKKSQ
HAKNSIFNVPEGIEETEGNKIATRI LKPSKGVKNPDLEATTPYISNLDSRLRLKGEVAKNWEVRQME LS
KVYQKQFERDAEHMIETLRPYKKACERSMGDVGGCANDVVRLSAEFEQFQIAVRGMEVSVKQVFDKENH
LRTIGARNQITDHVAARLAEWHLVKELMERRSAVLHNAKEFFTSAQRYFAEVPRTAQPQGVNPNDFLPM
QESLEEAIRKHDAFWAAQVEEVYAQAYDDASKVTRALKEADAEDNVAREHSSRLQRAHKQLMEKWKERQVL
LHHMLAMIAFETDVRLVVDWLEQHGEPYLRRNIHIGENVNQARSFQRNHTNFORVAANTYGNVQKLHQVY
QDVTKSGSKICDQVDTMHGLMTDLSAKIEKFTKIENSRELLLRQSVLFHTHYKELTDWYGMGEKYKDRRI
DLTVPTCDKNKERFVLETDETAQAYAMTIDEGKTLIEMMQRATRVFDVDSASVAHIQLLISDIEDKNTQ
ISSEWSPRRTLLHIIASKFAMFQNFQFVLEQIKSWEVDMREMAESDSFYDKADSVLPFHDMQOQVRNAT
ADIQSAKELSQALYSQKLTDLRDKRDRVVLEAIREHVRELEIAESRVMTYANETSMRMQAASEVGLRR
ISNSVCKLIDNQLSALTQLGVIPNDYNDTVQKQEE LRAFRDAVQHKLKEPYDAFVIRFRELME NRLANRD
EVVYHNEAIQAKYCRMLNMLCEDRNKLLKSAHGCKYKTYETAVLPILNQLESEYHSPITVDWCAGCTSSIDA
DRAAYVADLLSKHMDYKERFGKGCTYALRNGDFLLRYIRRSVTVQSERKHETKIADMKNNIRERQSNIL
ELWMQKLLLEGCQSFI FIEASAKELLEFMKGEGRNKLKHFEKKGRESTDDDDEFKFKSEVKQKTDI
QTFLMLSTNETMSRGVHGDEIDRCIEQVKDEF SRFSRRVGDCEVVLRGENGSTQSSKDEFSLDRHSDTAI
FNEKKINERREENRKMLEPMREL IQSERDYIKDLERCVMITVYKFDQAAKNGTITPLNPLKYEIFGNIEK
IFKFHMDKLLHELIKVENQPEAVGASFIVMIDLLNELYTEYCVNKEQKNHVIATPDVASFFTGIRERHGL
EINNEIASLLIKPVQRITRYRL LIEQLMRSCDTKDNDLKEAYEVVCSVPRKVMDLIHVNCIDLKDFKVDE
LGFVFTQDTLTFWEPRAYFKGRGKERQVFLFDISIVFAKRIEVS PKNIKYVIKGPPLPLSEVSI VEHVEG
DTCRFLRVGTVSSNDNRIDLKANNHHTKVKVQKIRDLTAGMLPLGLGVSEAYSVGTLSTARSVSVRS G
ASTSSGGENRQSDVESLLRHRYSVHSCDSEQSSEVWIVTADFDGQVEGHLTVHKGDRVEIVEDQATDCA
EYVQVVLCDQPTKHGLVPASIIAPPESGSPDRPDDSNASGSTVSKRKS LRRIFANSSKERA AAAANSSSS
NNMSPATRQSTSTSSPVTANGHTTESSTSAPIGVLSSGEPTSSNSIPSSSTPTHIVPSSVPATVPIVVA
EDEKAEDCLPPMENITINNSIDEKDMNEDSVAASDAVEPEVALPAKVEKTEETARFKRSYVLMELVET
EQDYVKDLTSVVEGYIGNLMKMDLPADLVGKDKLIFANIVNILEFHKTNPLKEIEKCSENYEAAGAA FVK
YERRLHTLYVTYQCNKPKSDYLLAQDDFEAFFADTKAKLGHKVALCDLLIKPVQRIMKYQLLLKDLILKFT
ERAKDKTDTLKKALQVMHVVPKACDDMMQVGR LQNFDKSLSAQKGLIHQGT LQISESIAGNVQKPKDRRI
FLFEQSALIADHIPPKKEFGMPTYIFKSQFMVNKMVFE PNVPDDPLRFVIKSSDPTQPTSFIANAQSQEE
KDEWNRKMSLELLDQQRKLLAALVDP RRYNDMS SGMGDL SLGAPGGDRRGAPSSANRSAASSSKKPAESP K
KEGKSKSSLFSFGKPKAKSPTSPPLDAAAAGKFQKVADDQVNLETD EKVKVLDVKNKYASVKKADGTVG
KCPSYFLTMSDIPGTNFAEQIQYRREWQRVDETDVEYGPSAFDTATSSSTDNFDV LLELTTCERPVVVED
MKDIEVVEGNDVEMCPIIISHTDFTVIWHGPAVD SKRARIQTNLNRLRIKHVKKCDAGAYSVIAKN SF
GVTSTVAFLSVISIDPPTDFTVKISGDHEVRLKWKAE SGLKYCIEYRILDDSSPENULIASTNIEKTHV
SLRNFARNSSYSFRVFAYNQVRVSAPSQCICICFDGTEM
```

NCBI Sequence Viewer - Netscape

File Edit View Go Communicator Help

Bookmarks Location: /query.fcgi?cmd=Retrieve&db=Nucleotide&list_uids=2944395&dopt=GenBank What's Related

SGMGDLISLGAPGGDFRGAPSSANRSAASSSKKPAESPKEGKSKSSLFSFGKKPARSP
TSPPPLDAAAAGKFQKVADDQVNLETDEKVKVLDVKNGYASVKKADGTVGKCPSYFLT
MSDIPGTNFAEQIQYRREWQRRVDETDVEYGPSAFDTATSSDNDVLELTTCERP
VVEDMKDIEVVEGMDVEMCPILSSHTDFTVIWHGPAVDSKRARIQTNQLNSRLLIKHV
KKCDAGAYSVIAKNSFGVTSTVAFLSVISIPDPPTDFTVKISGDHEVRLKWKAESGLK
YCIEYRILDDSSPENWLIASSTNIEKTHVSLRNFARNSSYSFRVFAYNQRRVSAPSQCIC
ICFDGTEM"

misc_feature 140..796
/gene="unc-73"
/note="N-terminal domain"

misc_feature 872..3538
/gene="unc-73"
/note="eight spectrin-like repeats"

misc_feature 3721..4306
/gene="unc-73"
/note="dbl homology (DH) domain"

misc_feature 4321..4663
/gene="unc-73"
/note="pleckstrin homology (PH) domain"

misc_feature 4835..5014
/gene="unc-73"
/note="SH3-like domain"

misc_feature 5216..5314
/gene="unc-73"
/note="PEST domain"

misc_feature 5534..6112
/gene="unc-73"
/note="dbl homology (DH) domain"

misc_feature 6128..6469
/gene="unc-73"
/note="pleckstrin homology (PH) domain"

misc_feature 7016..7288
/gene="unc-73"
/note="immunoglobulin domain"

misc_feature 7289..7576
/gene="unc-73"
/note="fibronectin type III domain"

3'UTR 7580..7611
/gene="unc-73"

polyA_signal 7593..7598
/gene="unc-73"

BASE COUNT 2385 a 1516 c 1865 g 1845 t
ORIGIN

Document: Done

NCBI Conserved Domain Database - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml> What's Related

NCBI

PubMed BLAST OMIM Taxonomy Entrez Structure

Search Entrez Structure for Go

CDD Help

Help on CD-Search and Databases

CD-Search

Search with advanced options

Smart

Explore the source database

Pfam

Explore the source database

MMDB

NCBI's structure database

Cn3D v3.0

3D-structure viewer

VAST

Structure comparisons

Research

CDD: A Conserved Domain Database and Search Service

Proteins often contain several modules or domains, each with a distinct evolutionary origin and function. The CD-Search service may be used to identify the conserved domains present in a protein sequence.

Computational biologists define conserved domains based on recurring sequence patterns or motifs. CDD currently contains domains derived from two popular collections, **Smart** and **Pfam**, plus contributions from colleagues at NCBI. The source databases also provide descriptions and links to citations. Since conserved domains correspond to compact structural units, CDs contain links to 3D-structure via **Cn3D** whenever possible.

To identify conserved domains in a protein sequence, the CD-Search service employs the reverse position-specific **BLAST** algorithm. The query sequence is compared to a position-specific score matrix prepared from the underlying conserved domain alignment. Hits may be displayed as a pairwise alignment of the query sequence with a representative domain sequence, or as a multiple alignment.

Run CD-Search:

Search Database:

Enter query as Protein

Read about **FASTA** format description, click [here](#) for advanced options.

Simple Keyword-Search for CDs:

NCBI CD-Search - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> What's Related

NCBI CD-Search Entrez ?

Search the **Conserved Domain Database** with Reverse Position Specific BLAST

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Search Database: All

Enter query as **Protein** Sequence in FASTA format Submit Query Reset

Please read about [FASTA](#) format description

Advanced options for the BLAST server:

Expect 0.01 Filter ☒ Low complexity. Search mode: Multiple hits 1-pass

Output formatting options:

Display up to 50 hits with Extended Graphic Overview in Color Scheme 3

Document: Done

NCBI CD-Search - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> What's Related

NCBI CD-Search Entrez ?

Search the **Conserved Domain Database** with Reverse Position Specific BLAST

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Search Database:

Enter query as **Protein**

```
>gi|2944396|gb|AAC12931.1| guanine nucleotide exchange facto
MGDYDGMKMTKAEDILHVLRDGIAVLPGGRCRAGQAVIVCPSREQPVNQDNLNRNVFLYLE
GFLVVIDMRGKQTWTNVRHILKALSSIESSTVQVFIIKPEKFEKQKAQMSLGTWDFEV
IIDSSHLPKTVGGSYPYDHDEWLELRDLEKWIWNITEIMEKLESVRREICEGEQPVDTV
HAKNSIFNVPVEGIETEGNKIATRILKPSKGVKNPDLEATTPYISNLTDLSRLLLKGEVAK
KVYQQKQFERDAEHMIETLRPYKKACERSMGDVGGCANDVRLSAEFEQFQIAVRGMEVS
```

Please read about [FASTA](#) format description

Advanced options for the BLAST server:

Expect [Filter](#) ☒ Low complexity. Search [mode](#):

Output formatting options:

Display up to with in

Document: Done

CD-Search Results - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> What's Related

NCBI CD-Search Entrez ?

RPS-BLAST 2.0.13 [May-26-2000]

Query= local sequence: gi|2944396|gb|AAC12931.1| guanine nucleotide exchange factor UNC-73A [Caenorhabditis elegans] (2488 letters)

Database: oasis_sap.v1.00
2830 PSSMs; 604,881 total columns

... This CD alignment includes 3D structure. To display structure, download [Cn3D v3.00!](#)

Mouse-over boxes to display more information

Sequences producing significant alignments:

		Score (bits)	E value
gn Smart RhoGEF	Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPa...	113	1e-25
gn Smart RhoGEF	Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPa...	97.9	5e-21
gn Pfam pfam00621	RhoGEF, RhoGEF domain	82.3	3e-16
gn Pfam pfam00621	RhoGEF, RhoGEF domain	77.3	8e-15
gn Smart SEC14	Domain in homologues of a <i>S. cerevisiae</i> phosphatidylinositol t...	64.0	8e-11
gn Smart PH	Pleckstrin homology domain.; Domain commonly found in eukaryot...	48.0	5e-06
gn Smart PH	Pleckstrin homology domain.; Domain commonly found in eukaryot...	47.3	9e-06

gn|Smart|RhoGEF, Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases; Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases Also called Dbp-homologous (DH) domain. It appears that PH domains invariably occur C-terminal to RhoGEF/DH domains. Improved coverage

Entrez-PubMed - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=9529254&dopt=Abstract What's Related

NCBI National Library of Medicine PubMed

PubMed Nucleotide Protein Genome Structure PopSet

Search PubMed for Go Clear

Limits Preview/Index History Clipboard

Display Abstract Save Text Order Add to Clipboard

About Entrez

Entrez PubMed
Overview
Help | FAQ
New/Noteworthy

PubMed Services
Journal Browser
MeSH Browser
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Old PubMed

Related Resources
Order Documents
Grateful Med
Consumer Health
Clinical Alerts
ClinicalTrials.gov

Privacy Policy

1: *Cell* 1998 Mar 20;92(6):785-95 [Related Articles, Books, Protein, Nucleotide](#)

UNC-73 activates the Rac GTPase and is required for cell and growth cone migrations in *C. elegans*.

Steven R, Kubiseski TJ, Zheng H, Kulkarni S, Mancillas J, Ruiz Morales A, Hogue CW, Pawson T, Culotti J

Samuel Lunenfeld Research Institute of Mt. Sinai Hospital, Toronto, Ontario, Canada.

unc-73 is required for cell migrations and axon guidance in *C. elegans* and encodes overlapping isoforms of 283 and 189 kDa that are closely related to the vertebrate Trio and Kalirin proteins, respectively. UNC-73A contains, in order, eight spectrin-like repeats, a Dbl/Pleckstrin homology (DH/PH) element, an SH3-like domain, a second DH/PH element, an immunoglobulin domain, and a fibronectin type III domain. UNC-73B terminates just downstream of the SH3-like domain. The first DH/PH element specifically activates the Rac GTPase in vitro and stimulates actin polymerization when expressed in Rat2 cells. Both functions are eliminated by introducing the S1216F mutation of unc-73(rh40) into this DH domain. Our results suggest that UNC-73 acts cell autonomously in a protein complex to regulate actin dynamics during cell and growth cone migrations.

PMID: 9529254, UI: 98188103

Display Abstract Save Text Order Add to Clipboard

[Write to the Help Desk](#)

Document: Done

NCBI CD-Search - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> What's Related

NCBI **CD-Search** Entrez ?

Search the **Conserved Domain Database** with Reverse Position Specific BLAST

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Search Database:

Enter query as **Protein**

```
>gi|2944396|gb|AAC12931.1| guanine nucleotide exchange facto
MGDYDGKMTMKAEDILHVLRDGIAVLPGGRCRAGQAVIVCPSREQPVNQDNLRNVFLYLF
GFLVVIDMRGKQTWTNVRHILKALSSIESSSTVQVFIIKPEKFWKQKAQMSLGTWDFEV
IIDSSHLPKTVGGSYPYDHEWLELRDLLEKWIWNITEIMEKLESVRREICEGEQPDVT
HAKNSIFNVPVEGIETEGNKIATRILKPSKGVKNPDLEATTPYISNLTDRLRLKGEVAK
KVYQQKQFERDAEHMIETLRPYKKACERSMGDVGGCANDVVRLSAEFEQFQIAVRGMEVS
```

Please read about [FASTA](#) format description

Advanced options for the BLAST server:

Expect **Filter** ☒ Low complexity. **Search mode:**

Output formatting options:

Display **up to** **with** **in**

Document: Done

CD-Search Results - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> What's Related

NCBI CD-Search Entrez ?

RPS-BLAST 2.0.13 [May-26-2000]

Query= local sequence: gi|2944396|gb|AAC12931.1| guanine nucleotide exchange factor UNC-73A [Caenorhabditis elegans] (2488 letters)

Database: oasis_sap.v1.00
2830 PSSMs; 604,881 total columns

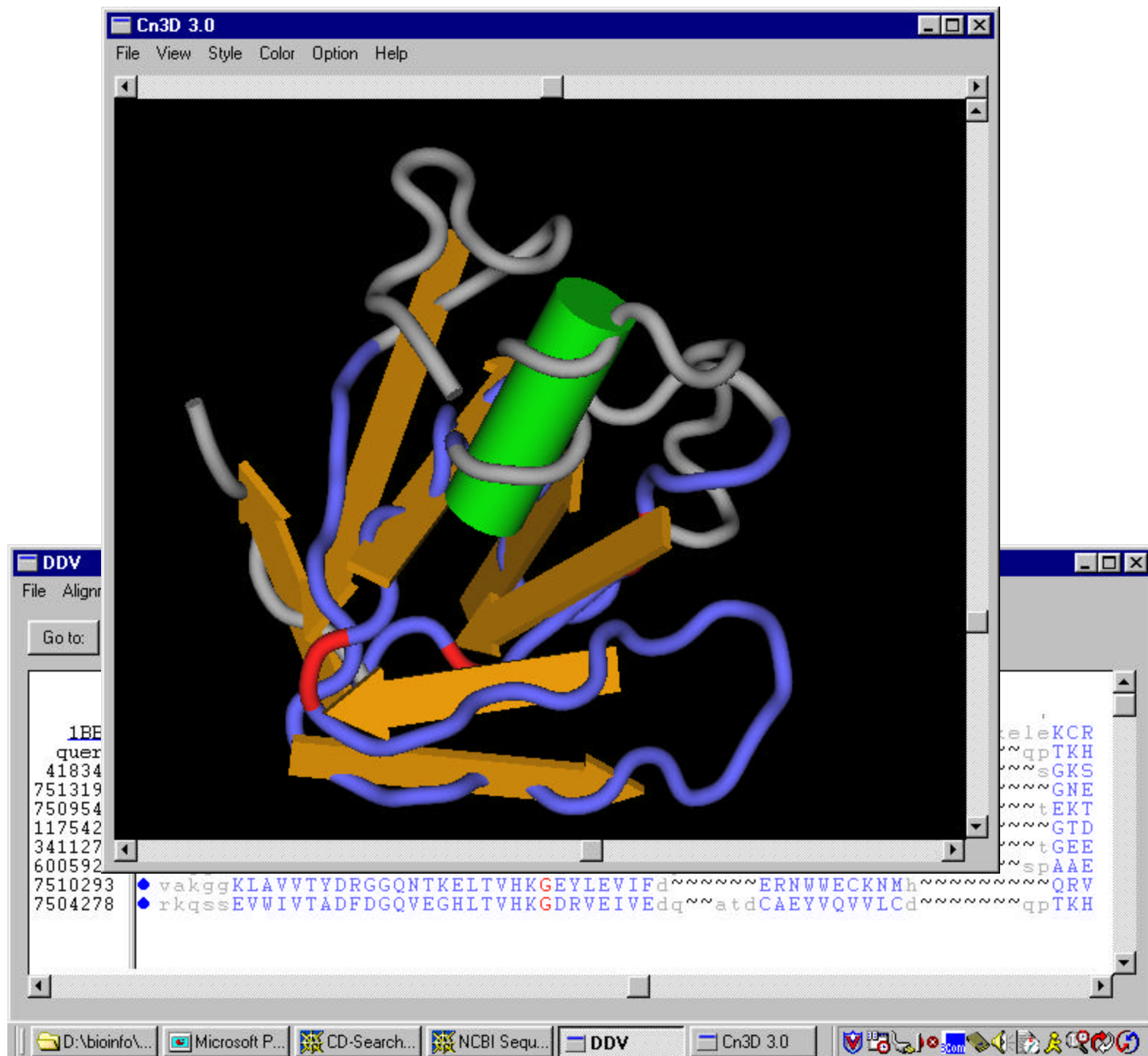
● .. This CD alignment includes 3D structure. To display structure, download [Cn3D v3.00!](#)

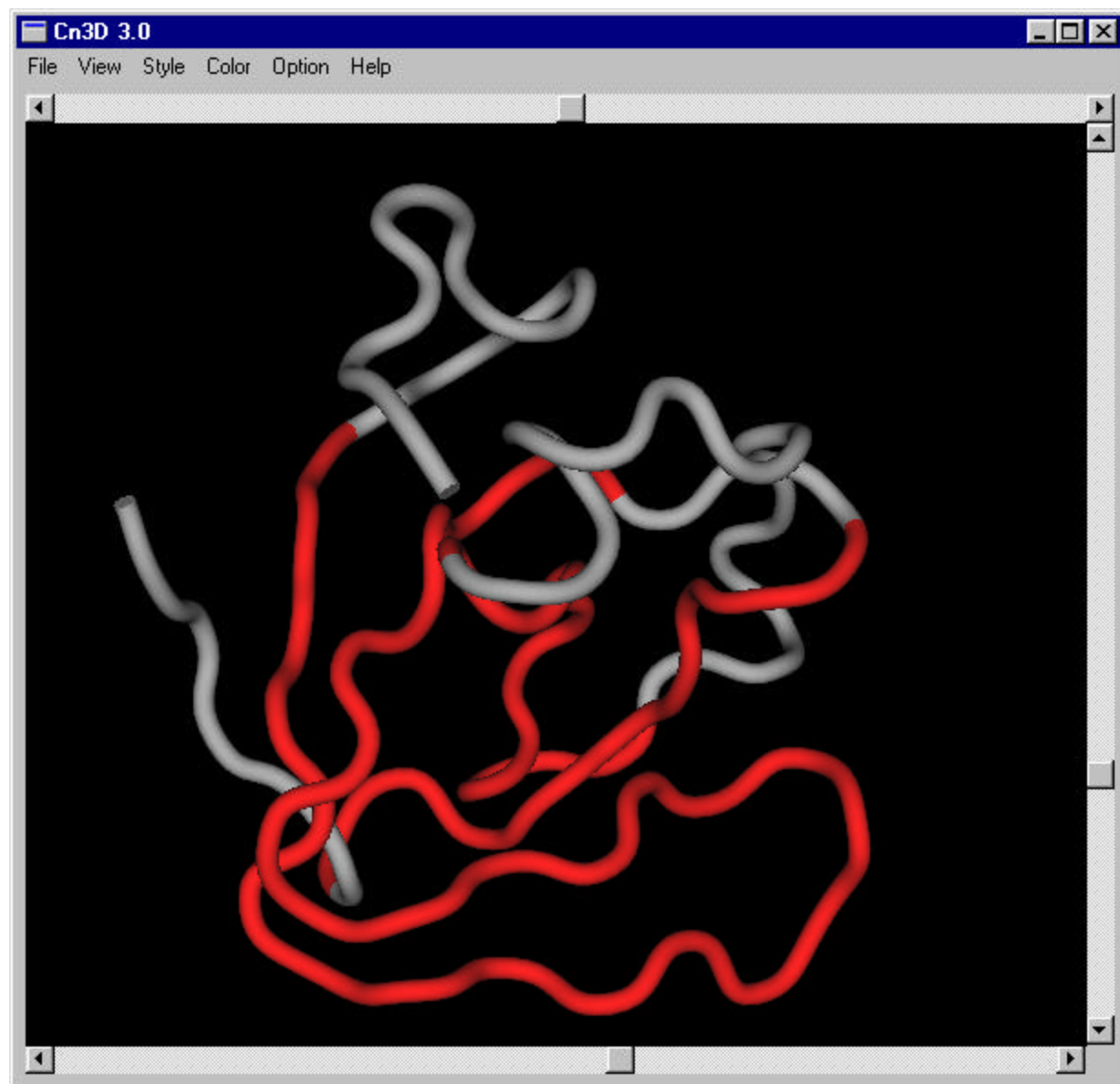
Mouse-over boxes to display more information

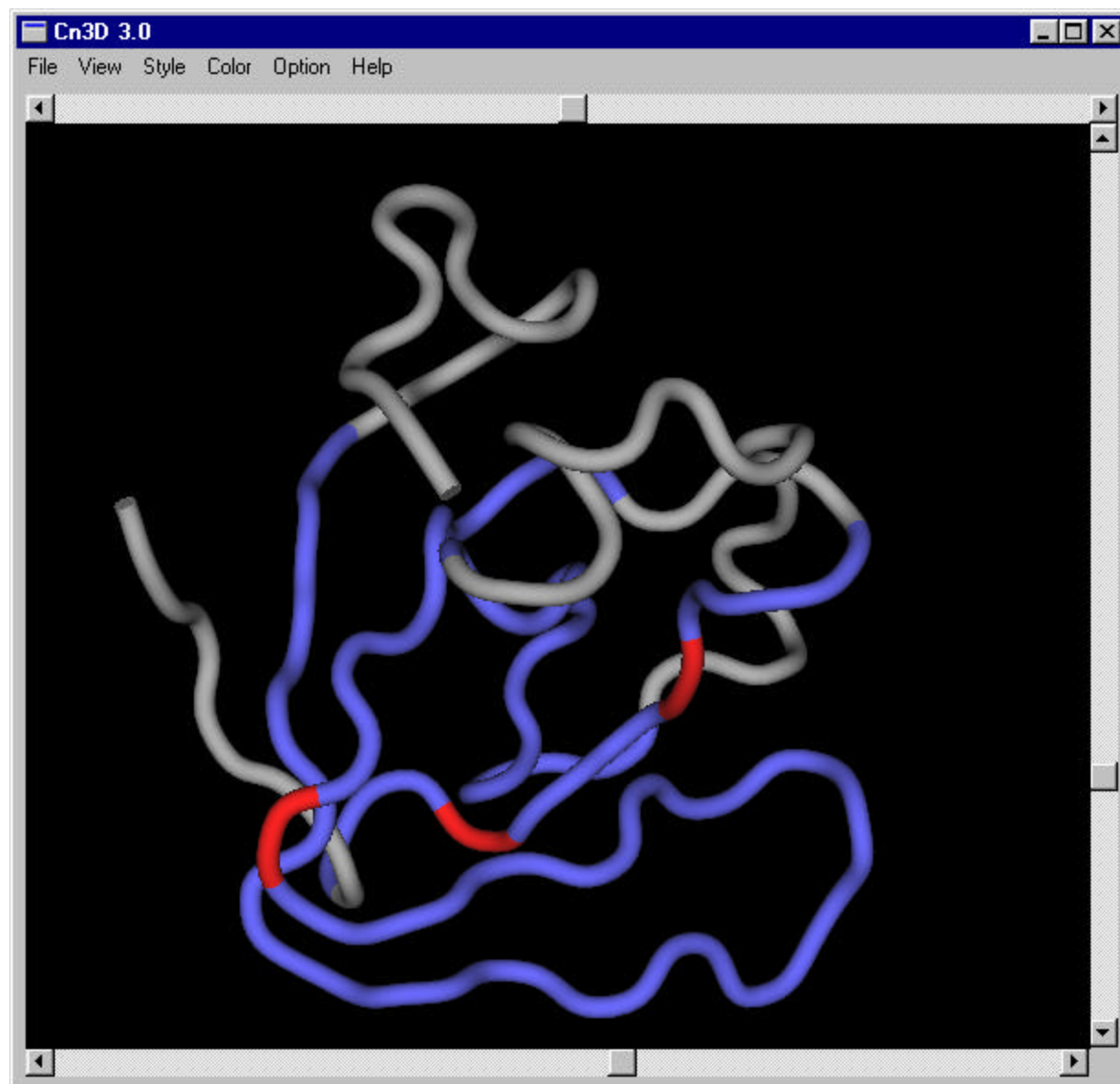
Sequences producing significant alignments:

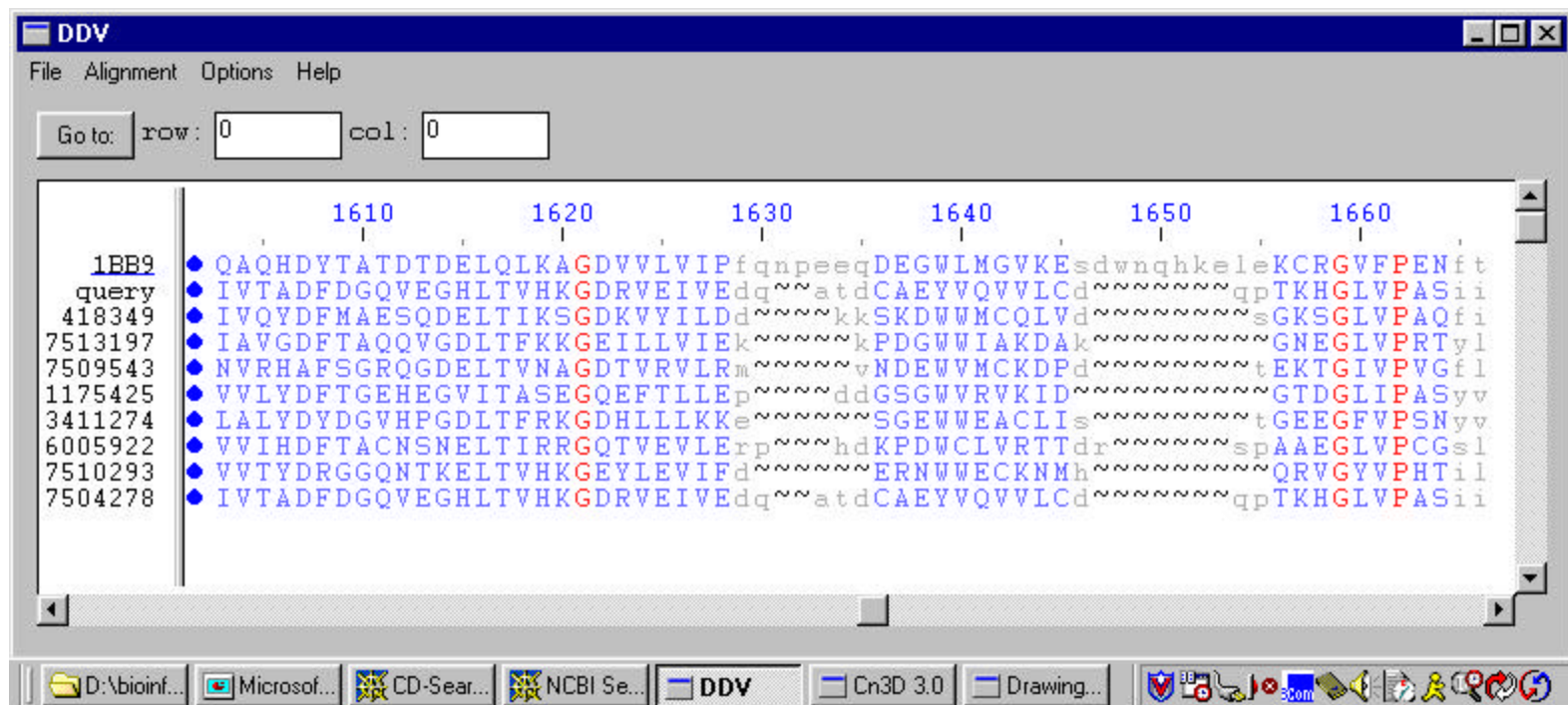
		Score (bits)	E value
●	gnl Smart RhoGEF Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPa...	113	1e-25
●	gnl Smart RhoGEF Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPa...	97.9	5e-21
	gnl Pfam pfam00621 RhoGEF, RhoGEF domain	82.3	3e-16
	gnl Pfam pfam00621 RhoGEF, RhoGEF domain	77.3	8e-15
●	gnl Smart SEC14 Domain in homologues of a <i>S. cerevisiae</i> phosphatidylinositol t...	64.0	8e-11
	gnl Smart PH Pleckstrin homology domain.; Domain commonly found in eukaryot...	48.0	5e-06
	gnl Smart PH Pleckstrin homology domain.; Domain commonly found in eukaryot...	47.3	9e-06
●	gnl Smart IG Immunoglobulin	37.1	0.010
●	gnl Smart SH3 Src homology 3 domains; Src homology 3 (SH3) domains bind to t...	36.0	0.023
●	gnl Pfam pfam00018 SH3, SH3 domain	34.0	0.087
●	gnl Pfam pfam00041 fn3, Fibronectin type III domain	32.8	0.20

Document: Done









CD-Search Results - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> What's Related

Subject: 61 TCKAVNSLG 69

gn|Smart|SH3, Src homology 3 domains; Src homology 3 (SH3) domains bind to target proteins through sequences containing proline and hydrophobic amino acids. Pro-containing polypeptides may bind to SH3 domains in 2 different binding orientations.

Add query to multiple alignment, display sequences

Length = 72
Score = 36.0 bits (81), Expect = 0.023

Query: 1575 EVWIVTADFDGQVEGHLTVHKGDRVEIVE--DQATDCAEYVQVVLCDQPTKHG 1625
Sbjct: 3 FKVQAQHDYTATDTDELQLKAGDVVLVIPFQNPPEEQDEGWLGMVKESDWNQHKLEKCRG 62

Query: 1626 LVPAS 1630
Sbjct: 63 VFPEN 67

gn|Pfam|pfam00018, SH3, SH3 domain

Add query to multiple alignment, display sequences

Length = 57
Score = 34.0 bits (76), Expect = 0.087

Query: 1582 DFDGQVEGHLTVHKGDRVEIVEDQATDCAEYVQVVLCDQPTKHGLVPASIIAP 1634
Sbjct: 8 NFHGTAEQDLPFCRGDVLITIVAVT--KDPNWKAKNKV--GREGIIPANYVQK 56

gn|Pfam|pfam00041, fn3, Fibronectin type III domain

Add query to multiple alignment, display sequences

Document: Done

CD multiple alignment for SH3 - Netscape

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi> What's Related

NCBI CD-Browser Entrez ?

CD: SH3, CD-Search result with query-sequence added

Description: Src homology 3 domains; Src homology 3 (SH3) domains bind to target proteins through sequences containing proline and hydrophobic amino acids. Pro-containing polypeptides may bind to SH3 domains in 2 different binding orientations.

Source: [Smart](#)

This CD alignment includes 3D structure. To display structure, download [Cn3D v3.0!](#)

Redisplay Alignment showing up to 10 sequences most similar to the query

☒ Aligned chains ☒ Virtual Bonds ☒ Launch Cn3D ☐ FASTA with gaps
☐ All chains ☐ All Atoms ☐ HTML Display ☐ Phylip format
☐ Text Display

		10	20	30	40	50	60	
1BB9	45	FKVQAQHDYTATDTDELQLKAGDVVLVI	PFQNP	EEQDEGWL	MGVKES	DWNQHKELEK	CRG	104
query	1575	EVWIVTADFDGQVEGHLTVHKGDRVEIVE	--DQATDCAEYVQVVLCD	-----	QPTKHG			1625
418349	356	KRGIVQYDFMAESQDELTIKSGDKVYILD	-----	SKDWWMCQLVD	-----	SGKSG		403
7513197	153	EEYIAVGDFTAQQVGDLTFFKKEILLVIE	-----	PDGWWIAKDAK	-----	GNEG		198
7509543	491	GFANVRHAFSGRQGDELTVNAGDTRVLR	-----	NDEWVMCKDPD	-----	TEKTG		537
1175425	518	VDGVVLYDFTGEHEGVITASEGQEFTLLE	-----	GSWVRVKID	-----	GTDG		563
3411274	54	VVLLALYDYGVPGLTFRKGDHLLK	-----	SGEWEACLI	-----	TGEEG		99
6005922	1600	ELTVVIHDFETACNSNELTIRRGQTVEVLERPHD	---	KPDWCLVRTTDR	-----	SPAAEG		1650
7510293	513	KLAVVTYDRGGQNTKELTVHKGLEYLIVIF	-----	ERNWWECKNMH	-----	QRVG		557
7504278	15	EVWIVTADFDGQVEGHLTVHKGDRVEIVEDQAT	--	DCAEYVQVVLCD	-----	QPTKHG		65

1BB9	105	VFPEN	109
query	1626	LVPAS	1630
418349	404	LVPAS	408

Document: Done

BIND - a database for molecular interactions

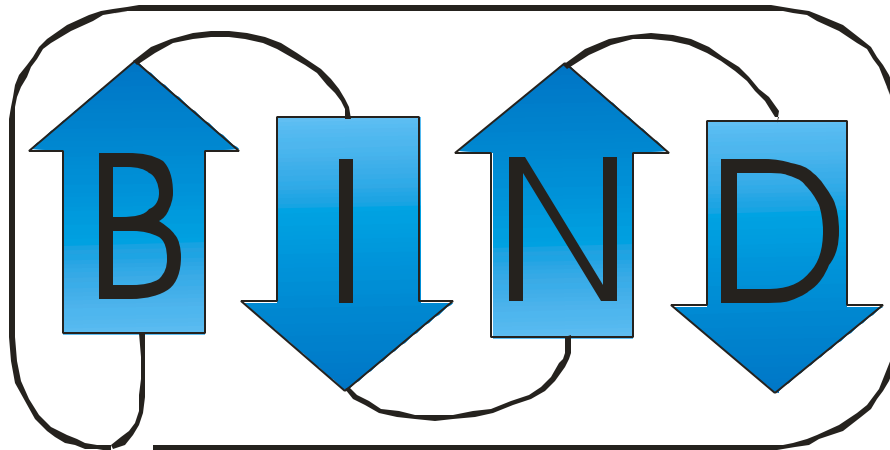
www.bindddb.org

Protein Assembly - A Database Problem

- How do we assemble all the molecules in the cell?
- How can we discover new pathways and functions?
- How do we simulate cellular processes?

An Interaction Database...

- Databases of molecular interactions are needed to understand the mechanistic complexity of diseases.
- BIND is the Biomolecular Interaction Network Database
 - molecules interacting with each other
 - captures the details of molecular function and mechanism



Interaction Databases

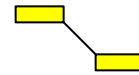


Interaction pair
"A binds B"

Database of Interactions

Molecule = Vertex

Interaction = Edge



Tools/Computations

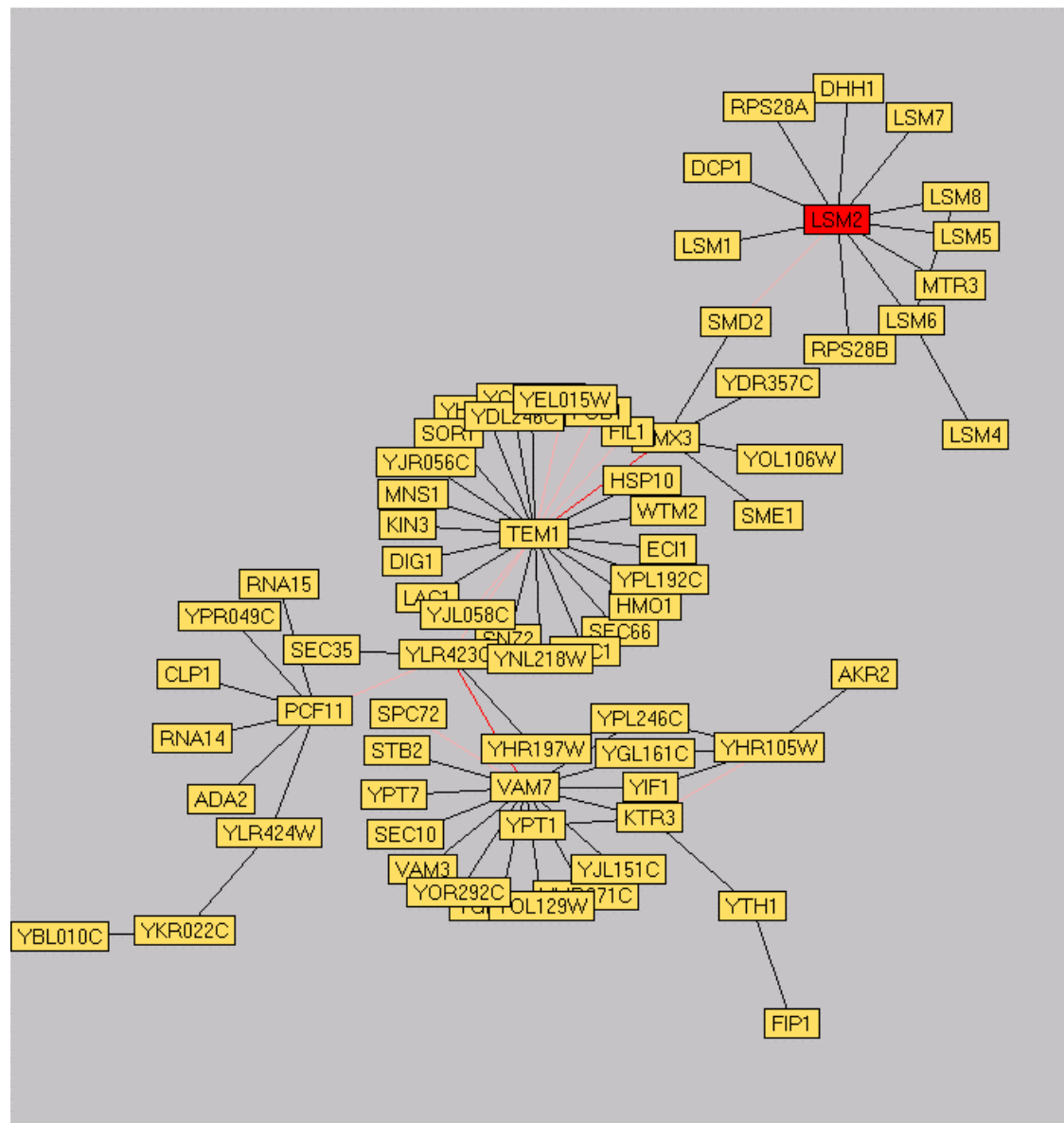
Graph Theory

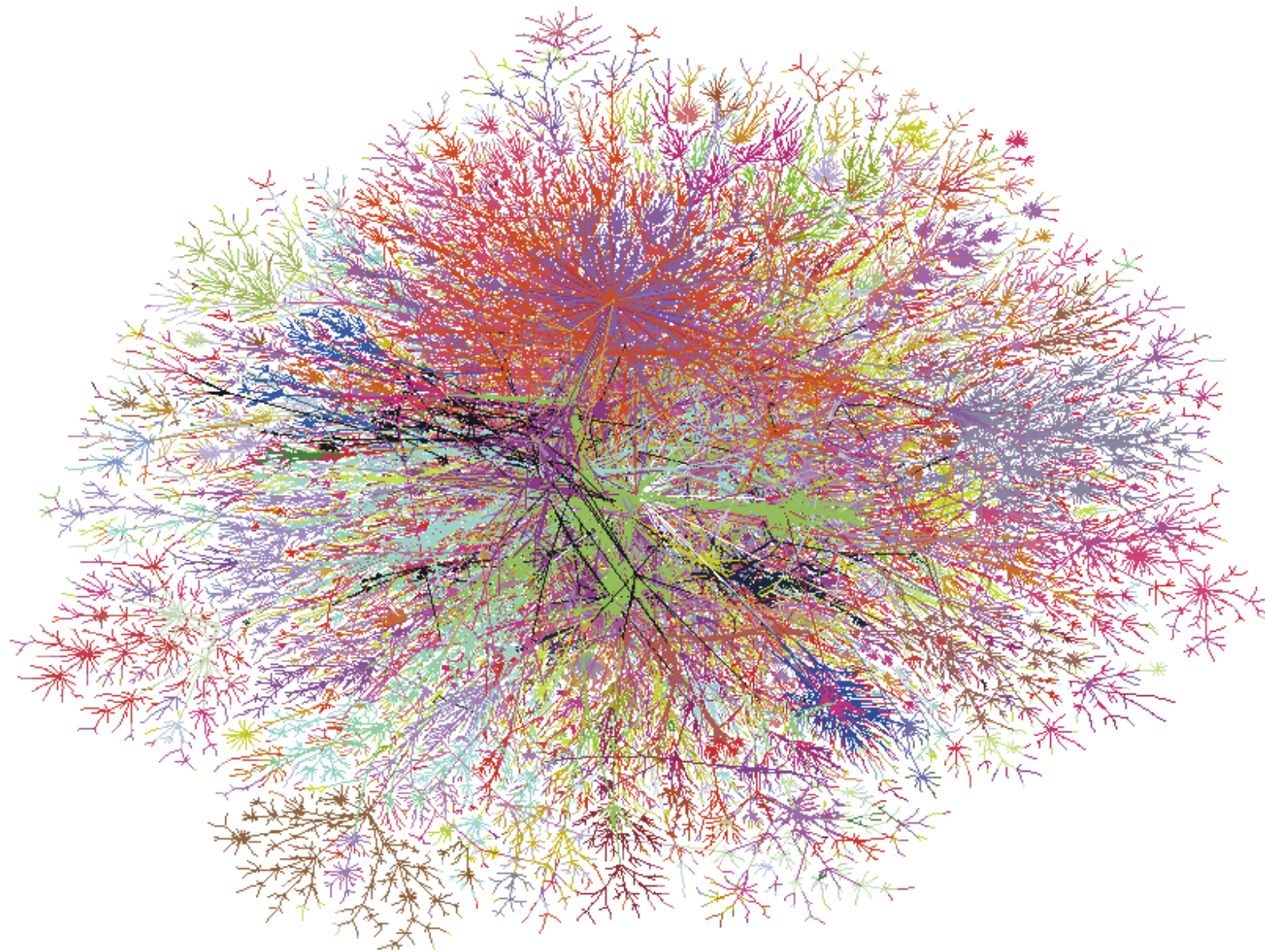
Pathway Finding

Simulations

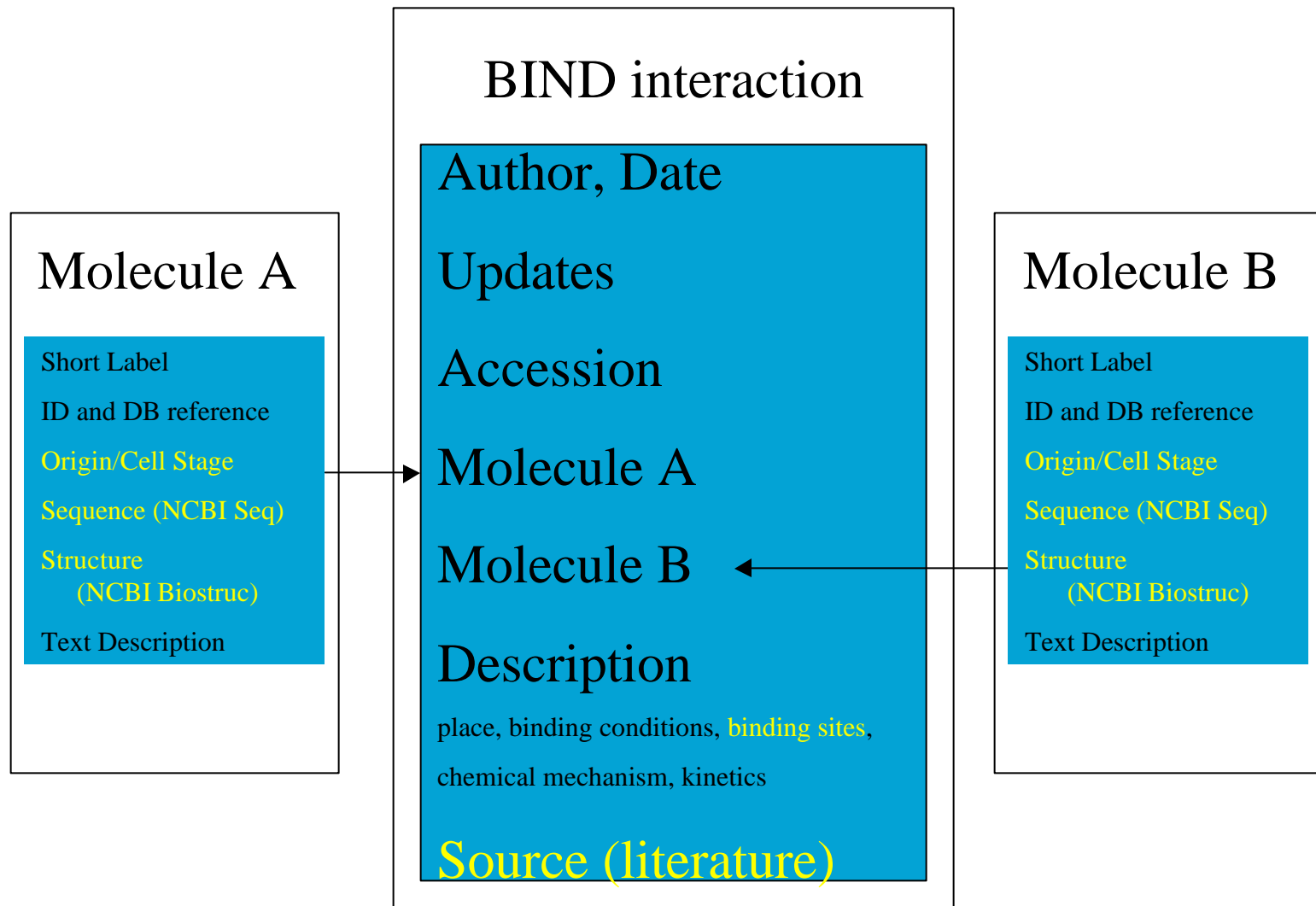
Cellular CAD

Y2H Data from Stan Fields in BIND





Standards are Essential



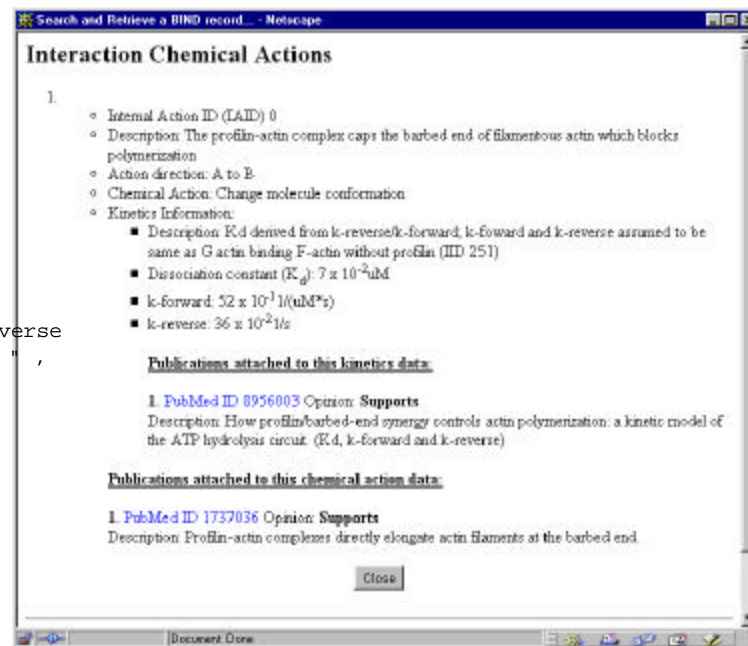
BIND is both ASN.1 and XML Compliant

XML

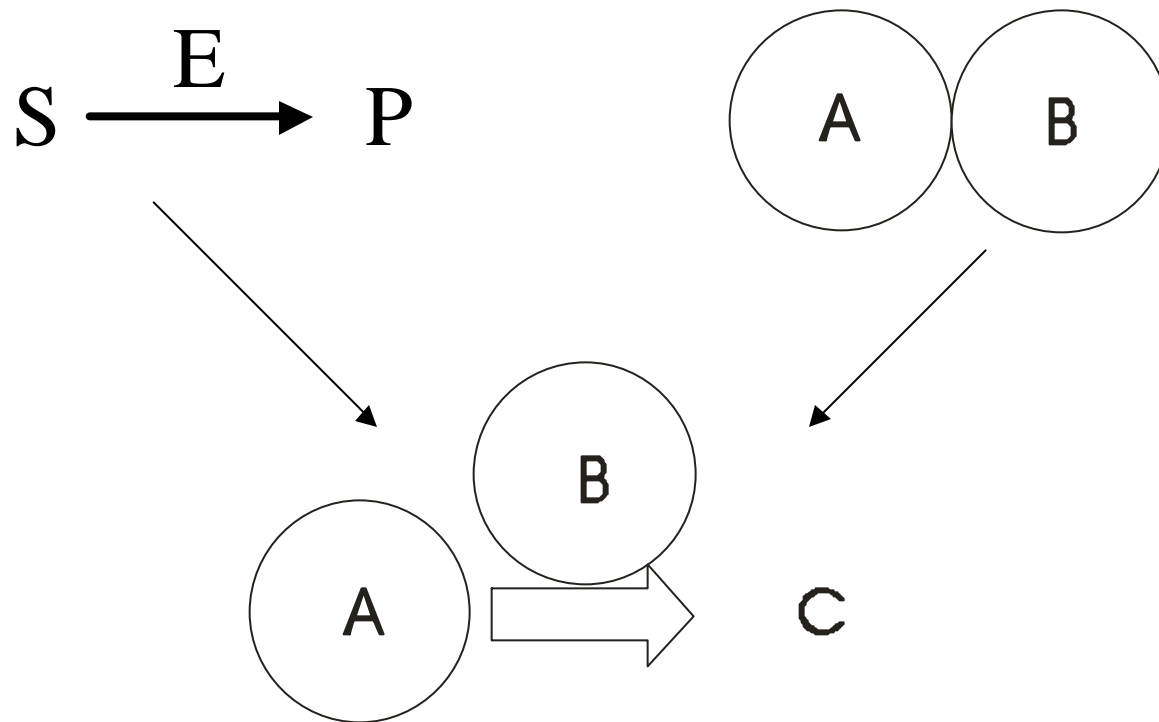
```
<BIND-action_kinetics>
  <BIND-kinetics>
    <BIND-kinetics_descr>Kd derived from k-reverse/k-forward; k-foward
and k-reverse assumed to be same as G actin binding F-actin without profilin (IID 251) </BIND-kinetics_descr>
    <BIND-kinetics_kd>
      <RealVal-Units>
        <RealVal-Units_scale-factor>-2</RealVal-Units_scale-factor>
        <RealVal-Units_scaled-integer-value>7</RealVal-Units_scaled-integer-value>
        <RealVal-Units_units>uM</RealVal-Units_units>
      </RealVal-Units>
    </BIND-kinetics_kd>
```

ASN.1

```
kinetics {
  descr "Kd derived from k-reverse/k-forward; k-foward and k-reverse
assumed to be same as G actin binding F-actin without profilin (IID 251) " ,
  kd {
    scale-factor -2 ,
    scaled-integer-value 7 ,
    units "uM" } ,
```

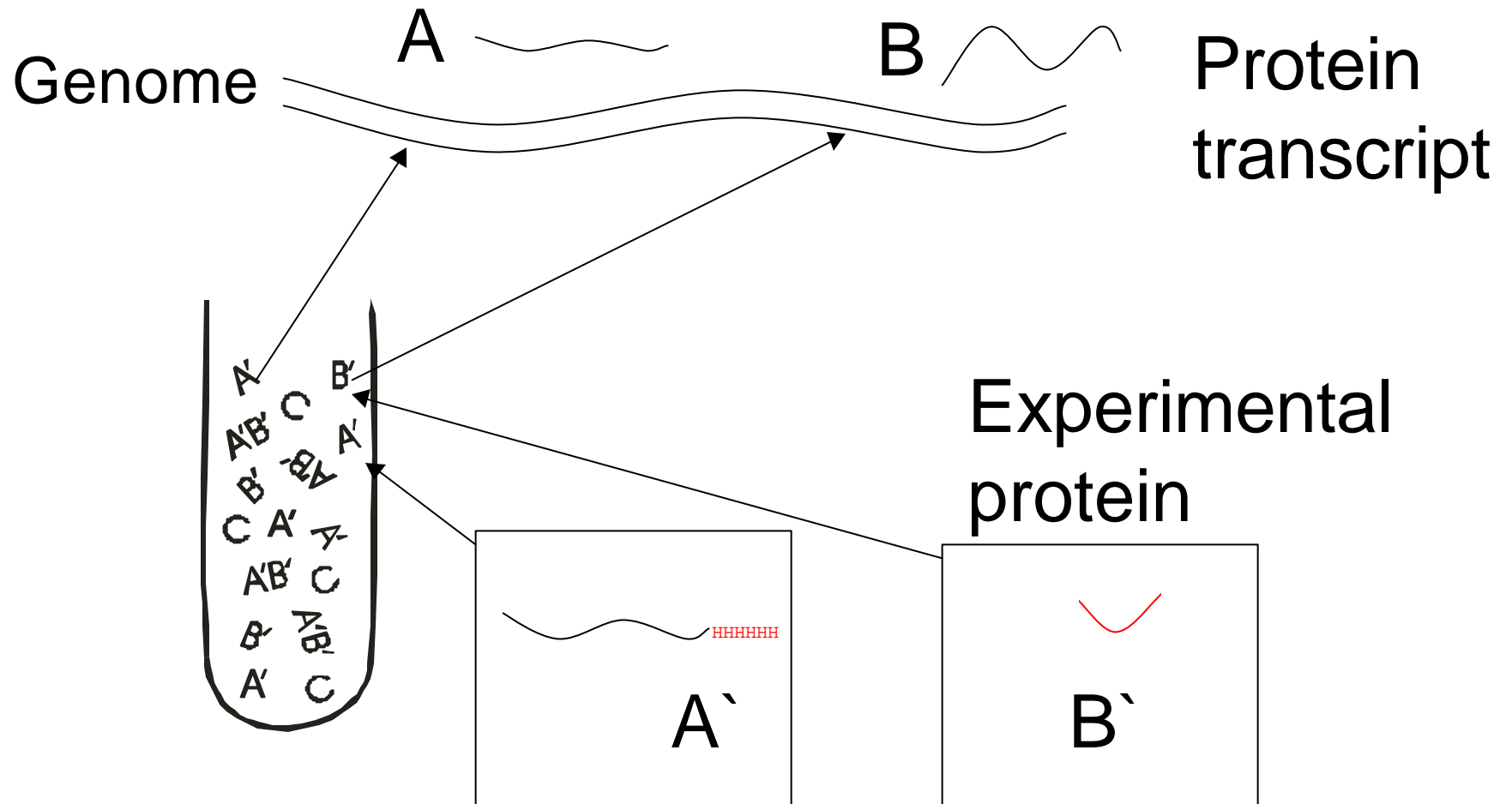


Reactions and Interactions

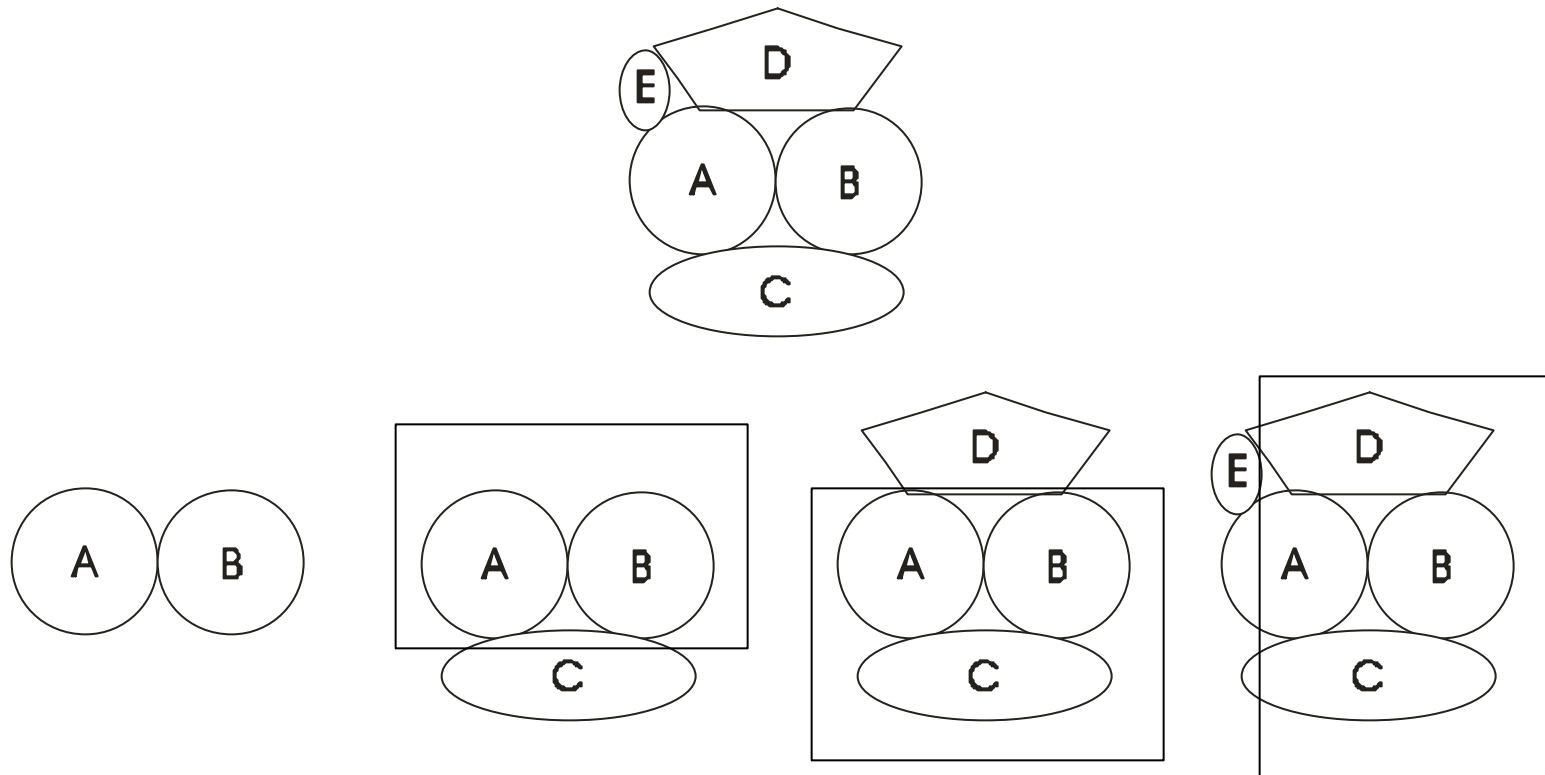


An interaction with a chemical change...

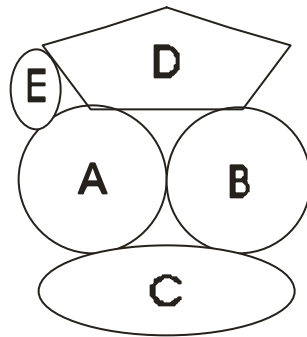
Linking Genomic and Proteomic data



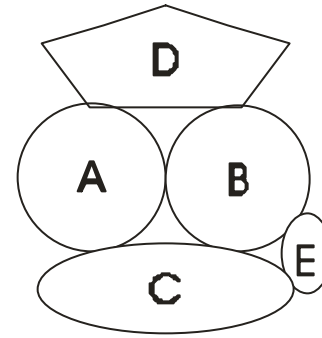
Complexes - Built up from Nested Interaction Pairs



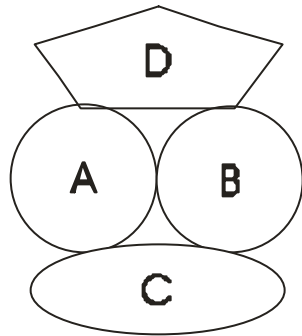
Uncharacterized Complexes



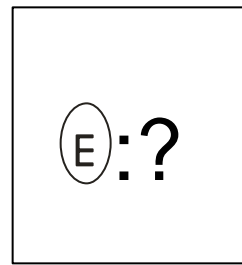
or



?



+



Immunoprecipitation
or affinity methods
discover complexes, not
always pairwise interaction

What BIND can encode...

- Simple interactions
- Enzymes, substrates and mechanisms including racemases and conformational changes
- Restriction enzymes
- Limited proteolysis (insulin, clotting cascade)
- Reversible phosphorylation
- Glycosylation
- Transcription factors
- Intron splicing
- Ubiquitin mediated protein degradation

What BIND cannot encode

- bulk phenomena
 - membrane potentials
 - gradients
 - calcium waves
 - water
- “perfect” cellular localization
 - (4-D time-development/organism axis).

File

Edit

View

Go

Communicator

Help

Data Manager

Menu

Version 1.5.1

About

Help

Search

Browse

Interactions

- Add
- Change

Pathways

- Add
- Change

Molecular Complexes

- Add
- Change

BIND Statistics

Administration

BIND stands for the Biomolecular Interaction Network Database.

NEW

Browse the **BIND v2.0 ASN.1 Specification**
The BIND specification has been published.
Get **help** on BIND

NEW

BIND Tutorials using pathway examples

INAD Pathway

EGF Receptor Pathway

High-throughput methods for identifying and characterizing all of the proteins, protein domains and protein interactions in a cell will eventually create more recorded biological information than the Human Genome Project.

Each protein expressed in a cell can interact with various different proteins and other molecules in the course of its function. Protein-protein interactions are often mediated by modular protein domains. One example is the SH3 domain which binds a proline rich motif. These "interaction networks" form conventional signaling cascades, transcription activation complexes, vesicle controlling mechanisms, cellular growth and differentiation systems, among other cellular machinery.

We have designed and implemented a new database encompassing the growing network of protein and other biomolecular interactions, called **BIND** (**B**iomolecular **I**nteraction **N**etwork **D**atabase).

This database will span the complexity of interaction information gathered through experimental studies of biomolecular

Document: Done

File Edit View Go Communicator Help

Interaction ID: 13
Accession date: Aug 18,
Description: The T cell ac
[View record update histo](#)

Data Manager
Menu

Version 1.5.0 Beta

About Help
Search Browse

Interactions

- Add
- Change

Pathways

- Add
- Change

Molecular Complexes

- Add
- Change

BIND Statistics

Administration

Molecule A

SLP-76
Description: SH2 domain
SH2 and SH3 binding mo
Molecule Type: Protein
GI: 6754518 ([NCBI](#)) ([B](#))
Molecule origin: Organism
Organism: [Mus musculus](#)

Molecule B

Vav
Description: proto-oncog
Molecule Type: Protein
GI: 6755955 ([NCBI](#)) ([B](#))
Molecule origin: Organism
Organism: [Mus musculus](#)

Visualize Interaction!

View other

Main Info	Publications	ASN.1
Cellular Place	Experimental Condition	Conserved Sequence
Cellular Place	Experimental Conditions	N/A
Binding Sites	Chemical action	Chemical State

NCBI Sequence Viewer - Netscape

NCBI

PubMed Nucleotide Protein Genome Structure

Search

Protein

 for

Limits Index History Clipboa

Display

FASTA

 View as

HTML

 Save Add to Clipboard

☐ Hide Brief and LinkBar

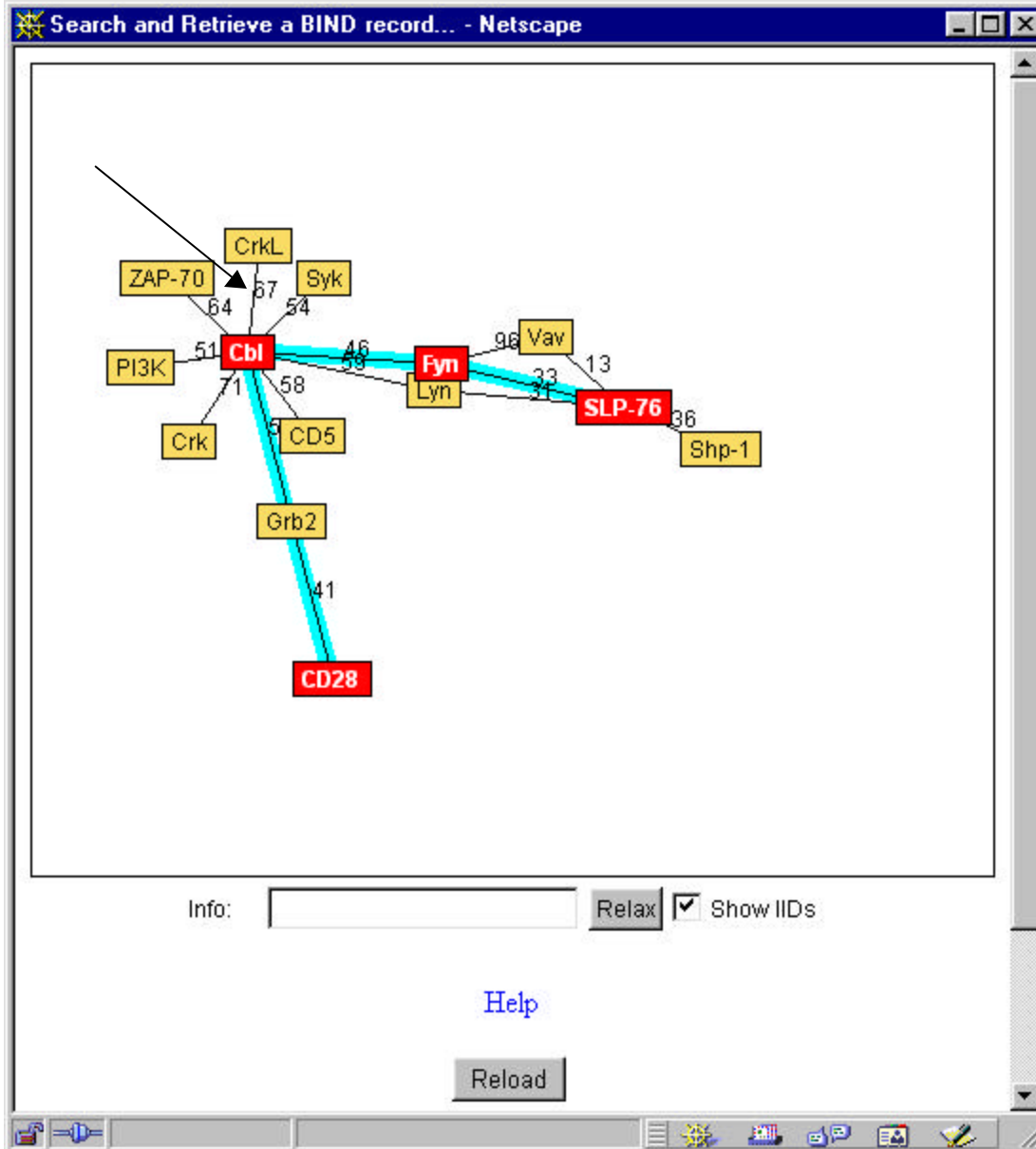
☐ 1 : GI = "6754518" [FASTA] lymphocyte cytosolic protei... [Related Articles](#), [Protei](#)

>gi|6754518|ref|NP_034826.1| lymphocyte cytosolic prote
MALKNVPFRSEVLAWNNDLADYFRKLNRYDCEKAVKKYHIDGARFLNLTENDI
DINKNEERRSIFTRKPKIPRFLEETESHEEDDGGWSSFEDDYESPNDDDPDGED
DAADYEPPPSNNEEALQSSILPPNSFHNTNSMYIDRPPTGKVSQQPPVPLRPKI
PPHPNHEEPSRSGNNKTAKLPAPSIDRSTKPPLDRSLAPLDREPFILGKKPPFSI
KPPLPPAMDRHERNERLGPVTRKPSVPRHGRGPDREDDDVHQRPLPQPSLI
SKNTFPLAHMPGAFSESNIGFQQSASLPPYFSQGPGNRPPLRSEGRNLPLVPNF
EWYVSYITRPEAEALRKINQDGTFLVRDSSKKTANNPYVLMVLYKDKVYNIQIF
GKEDFLSVSDIIDYFRKMPLLLIDGKNRGSRYQCTLTAAAGCL

BIND Data Manager - Netscape

File Edit View Go Communicator Help

Data Manager Menu Interaction ID: 67



Cbl (Protein, Mus musculus) and **CrkL** (Protein, Mus musculus).

Description: The proto-oncogene Cbl interacts with the adaptor protein, CrkL

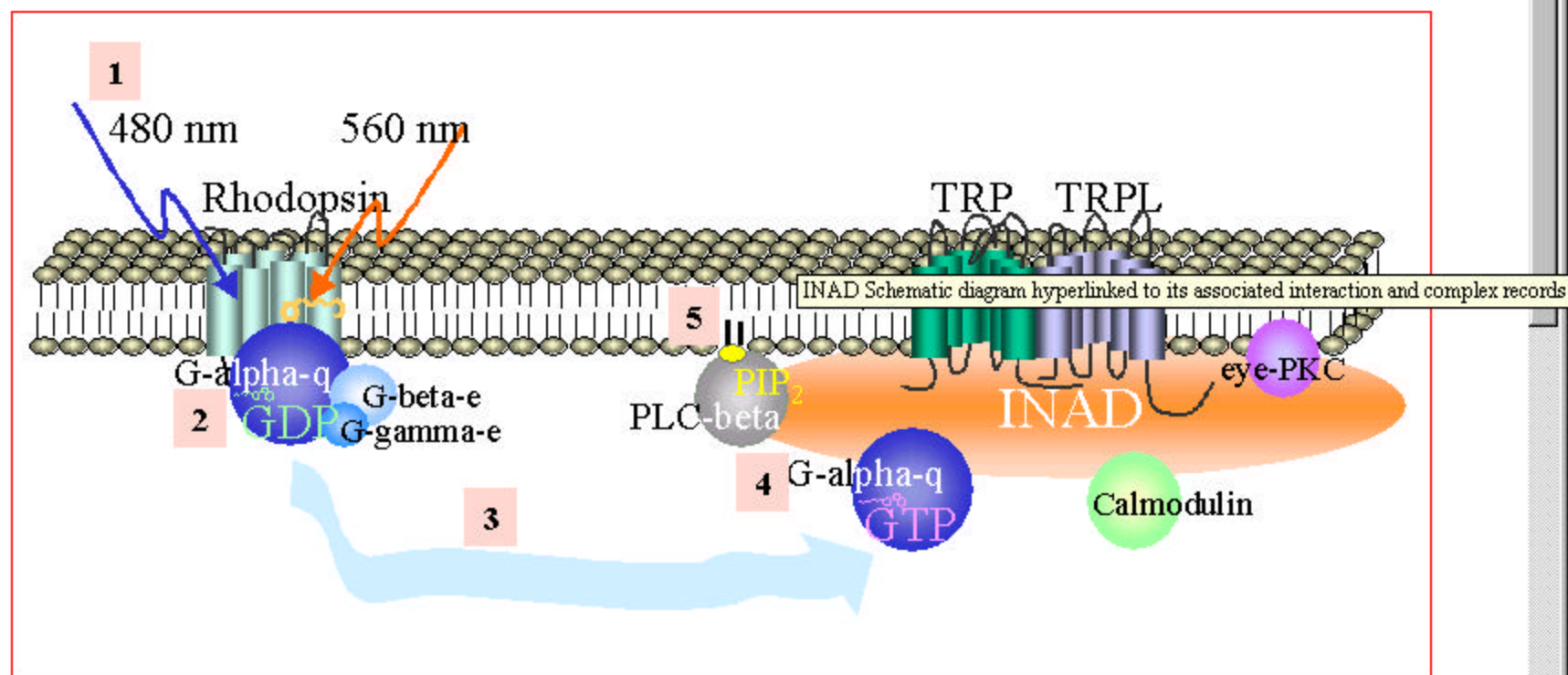
[\[Full Record\]](#) [\[2 PubMed abstracts\]](#) [\[2 BIND Publication links\]](#)

[\[View ASN.1 Report\]](#)

BIND Accession Number Query

Done

The INAD Pathway in Drosophila Photoreceptors - A Tutorial



The molecular events involved in visual transduction in fruit flies (*Drosophila melanogaster*) have been well studied. Specialized regions of the photoreceptor cells contain the components for receiving light signals. The INAD pathway translates a photon into a biochemical signal:

- 1 A photon of 480 nm is absorbed by rhodopsin resulting in the generation of metarhodopsin.

The INAD Pathway in Drosophila Photoreceptors - A Tutorial

Describing the complex anchored by INAD

185



Diagram of IID #185

IID #185

This record and the following ones represent pairwise interactions, i.e., simple examples of "A binds B".

[Want more info? Click here](#)

IID #186

This record is an example of a [modeled](#) interaction. The investigators examined the interaction between Drosophila INAD and Bovine Calmodulin assuming that it models the interaction that occurs in Drosophila. Note the statements about the modeled interaction in the [Interaction Description](#) and [Molecule B description](#).

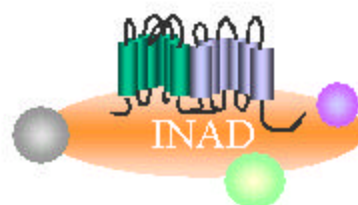
186



MCID 24

IID #185 + IID #186 + IID #187 + IID #188 + IID #189

=



INAD complex

Describing the G protein heterotrimer

203



G-β-e

+



G-γ-e



G-β-γ-e

MCID #24

This is a molecular complex record. It is generated by combining the above five interaction records. It contains [Publication](#) references that specifically discuss the existence of this complex.

[More](#)

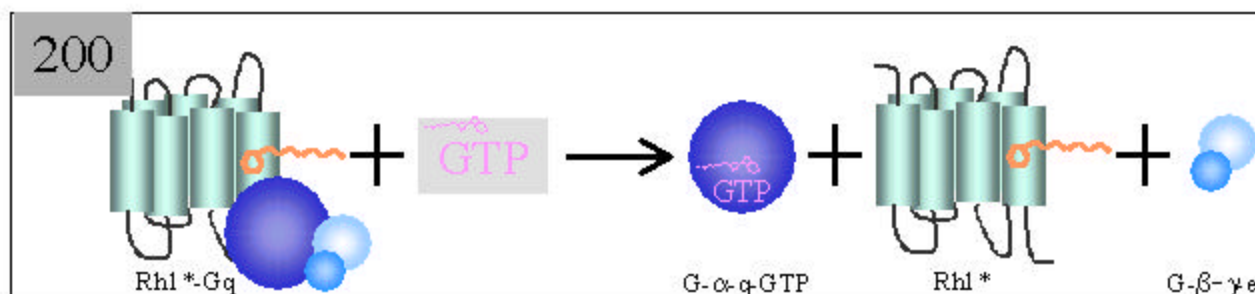
IID #203

These three interaction records represent the sequence of events that occur in order to assemble the heterotrimeric G protein complex, Gq. This one interaction record is used in Molecular Complex record #22 to describe the dimeric complex of G-beta-e and

G-alpha-q is open.

[More](#)

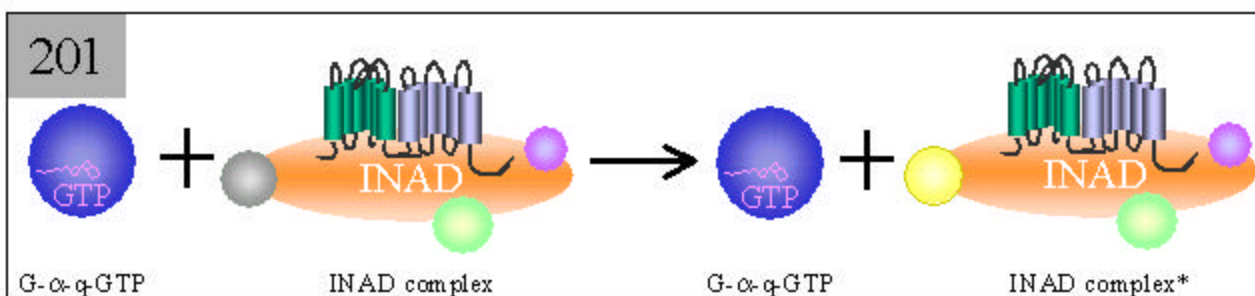
IID #200



This interaction describes the binding of GTP by G-alpha-q in the "open" conformation (as part of a complex). In order for this interaction to occur, Molecule A has a required **Chemical State** that was generated by the **Chemical Action** in IID #199.

[More](#)

IID #201



In this record, G-alpha-q-GTP interacts with the INAD complex and activates the enzyme PLC-beta. Here again, Molecule A has a required **Chemical State** (**active**) which leads to its **Chemical Action** on Molecule B (**change conformation**) to activate

Extended Description of IID #202 - Netscape

File Edit View Go Communicator Help

IID #202

The next step in this representation involves the active PLC-beta enzyme acting on its substrate PIP2 (phosphatidyl inositol-4,5-bisphosphate).

This is represented under **Chemical Action** as **remove** DAG (diacylglycerol) from PIP2 leaving IP3 (inositol-1,4,5-trisphosphate). These molecules will later be described under the **Chemical Result** field. When **add** or **remove** is selected as the Chemical Action, the object which is added or removed must be specified as an object from another database (e.g. GenBank, LIGAND, Klotho).

Eventually BIND will have its own database of small molecules from which objects can be selected for interaction records.

Back to tutorial with Interaction Diagrams

Comments and suggestions to:
wolting@mshri.on.ca

Interaction

Interaction ID: 202
Accession date: Apr 4, 2000
Description: Active phospholipase-C-beta as part of the INAD complex cleaves phosphatidyl inositol-4,5-bisphosphate to inositol-1,4,5-trisphosphate and diacylglycerol.
[View record update history](#)

Molecule A

INAD complex

Description: The multiprotein signalling complex anchored by INAD involved in Drosophila photoreceptor signal transduction
Molecule Type: Molecular Complex
Molecular Complex id: 24
Molecule origin: Organismal
Organism: [Drosophila melanogaster](#)

Molecule B

PIP2

Description: Phosphatidyl inositol-4,5-bisphosphate. Aliases: 1-Phosphatidyl-D-myo-inositol 4,5-bisphosphate
Molecule Type: Small Molecule
Database name: LIGAND
LIGAND Text ID: C04637
Molecule origin: Organismal
Organism: [Drosophila melanogaster](#)

Visualize Interaction!

View other information?

Main Info

Publications

ASN.1

XML

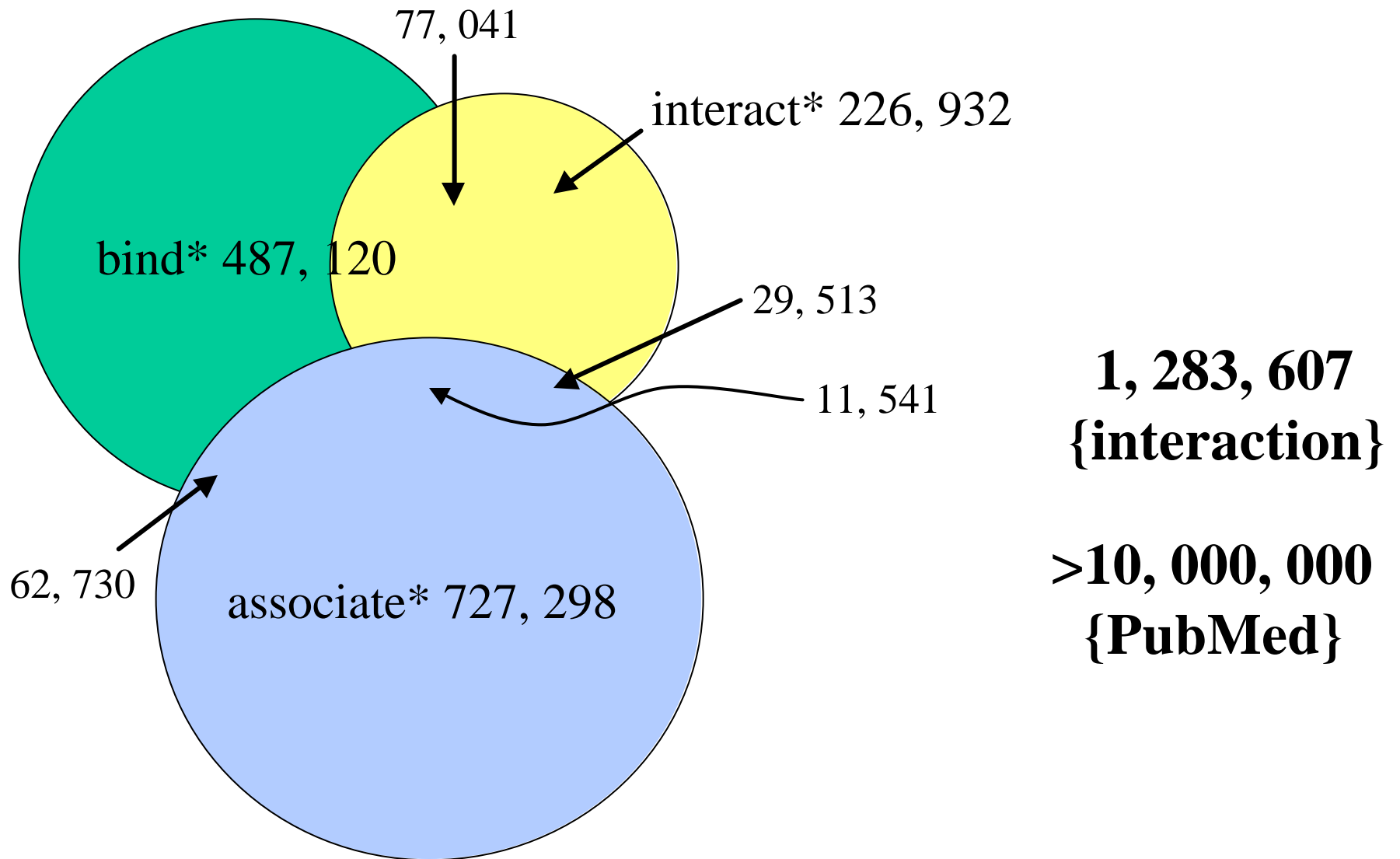
Cellular Place

Experimental Condition

Conserved Sequence

Document: Done

The backfilling problem



Identifying interaction papers

Vacuole SNAREs, including the t-SNAREs Vam3p and Vam7p and the v-SNARE Nyv1p, are found in a multisubunit "cis" complex on isolated organelles.

A major educational effort is needed to inform patients, providers, and third parties that the presence of a particular allele in an individual is only one piece in a complex puzzle of environmental and genetic interactions that may ultimately result in disease.

PreBIND: the backfilling solution

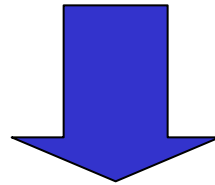
Create tools that help a few people enter records quickly and simultaneously offer these tools to the community as something that is immediately useful.

PreBIND is a data-mining tool which locates interaction references for a protein.

PubMed name search



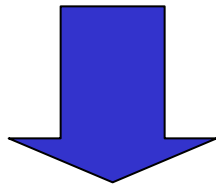
RPB2, Soh2, Rpo22



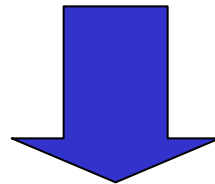
RPB2 [WORD] AND “Saccharomyces cerevisiae” [MESH]

Soh2 [WORD] AND “Saccharomyces cerevisiae” [MESH]

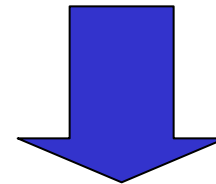
Rpo22 [WORD] AND “Saccharomyces cerevisiae” [MESH]



PMID's



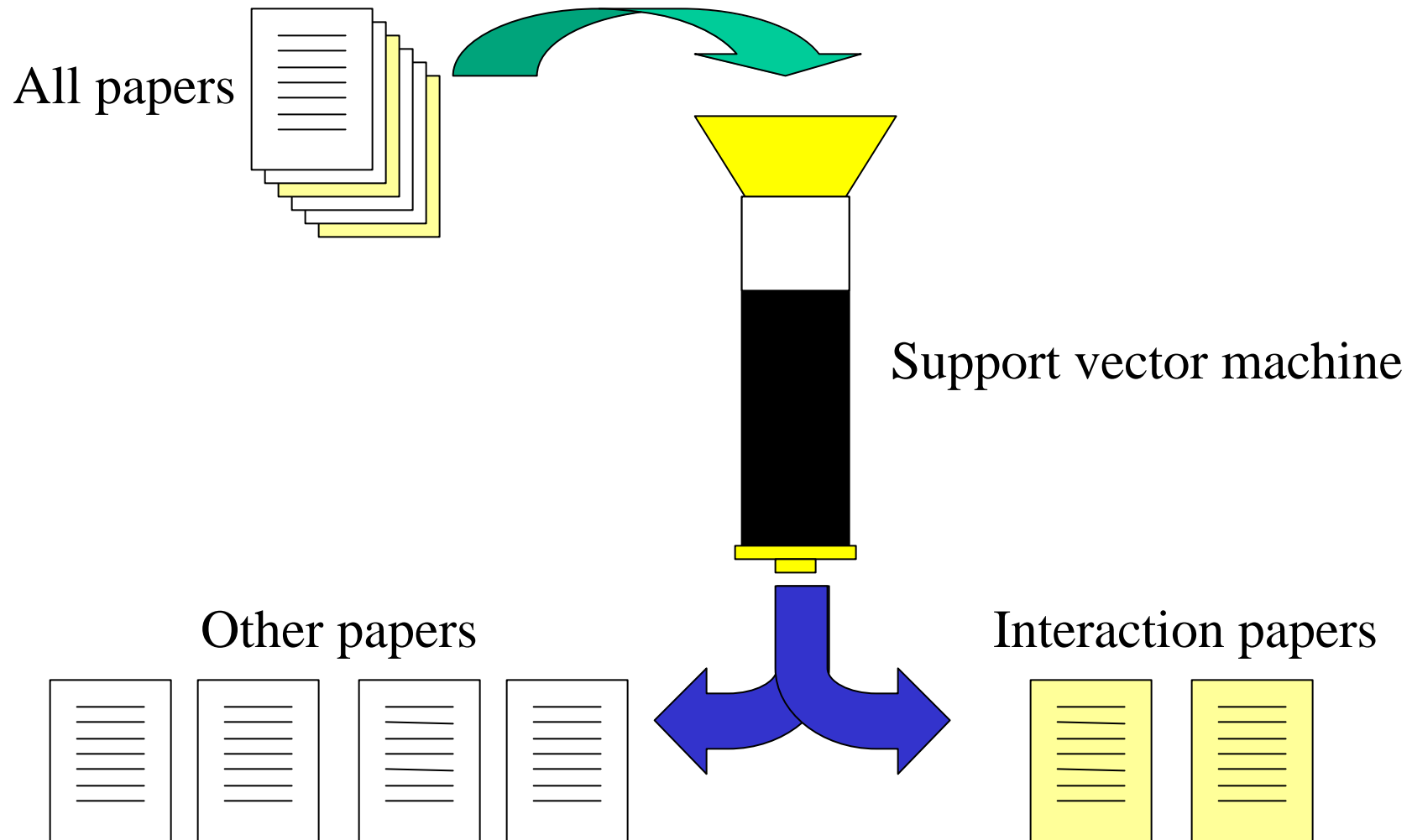
PMID's



PMID's

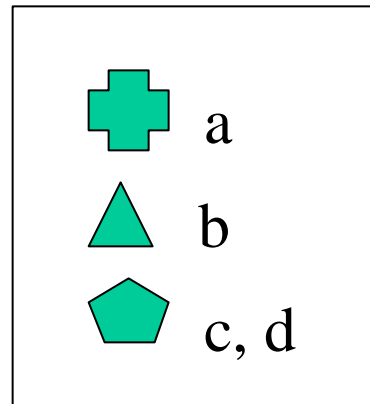
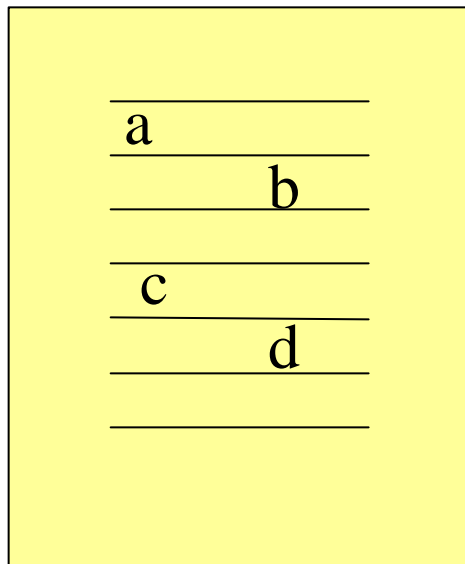


Identifying interaction papers



Identifying potential interactions

1. Classified as interaction paper
2. Talks about two or more proteins



a	and	b
b	and	c/d
a	and	c/d

Which interaction(s) does the paper talk about?

Scoring potential interactions

Soh1 Soh2

Soh1
Rad5p



Soh1



Soh2



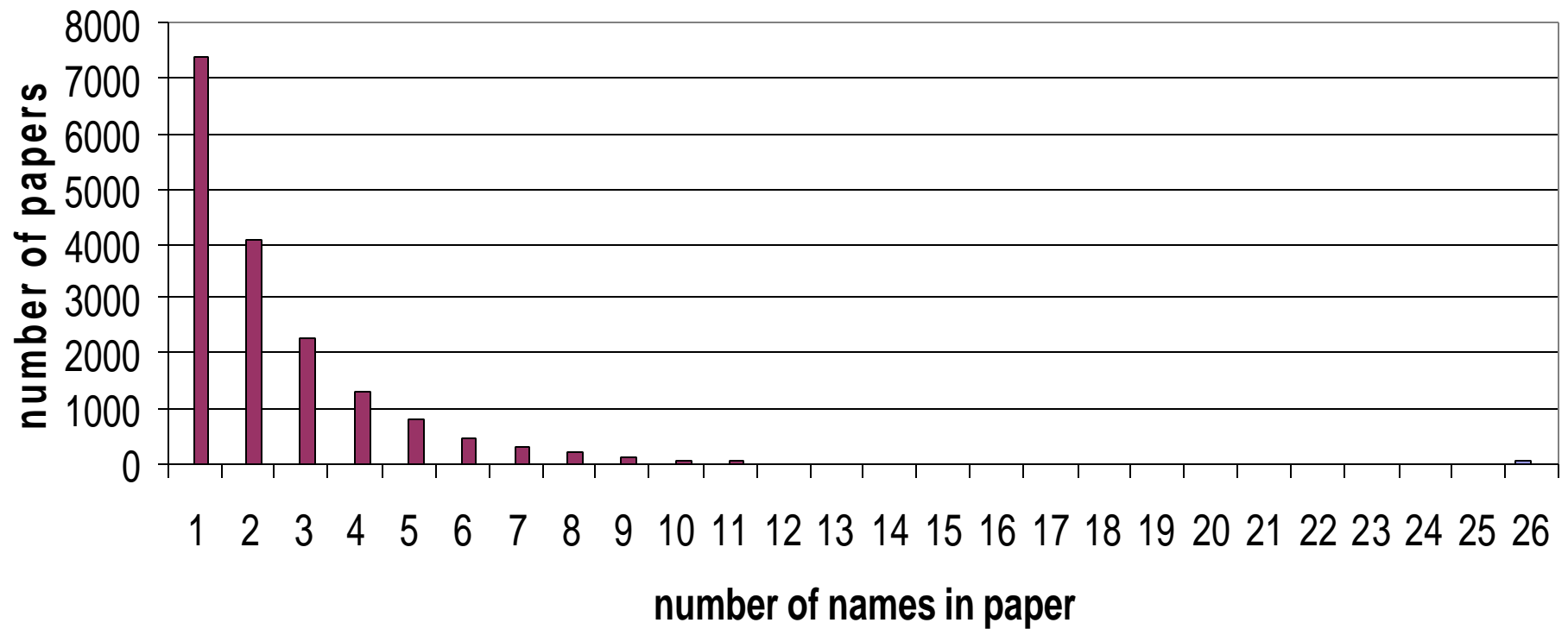
Rad5p

Soh1 and Soh2
Soh2 and Rad5p
Soh1 and Rad5p

The soh1, soh2 and soh4 mutants were isolated as suppressors.....

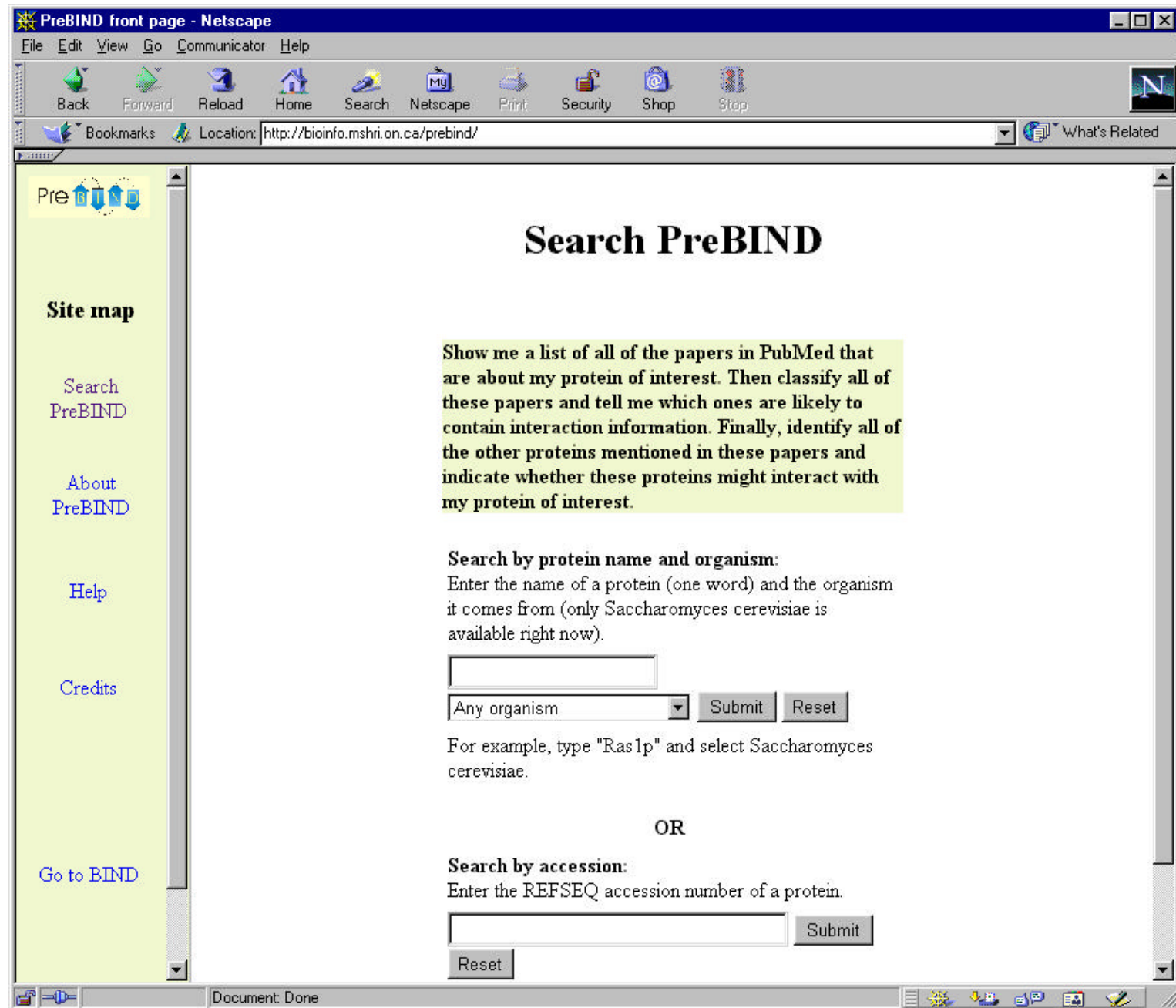
The Soh1 protein interacts with a DNA repair protein, Rad5p, in a two-hybrid system assay.

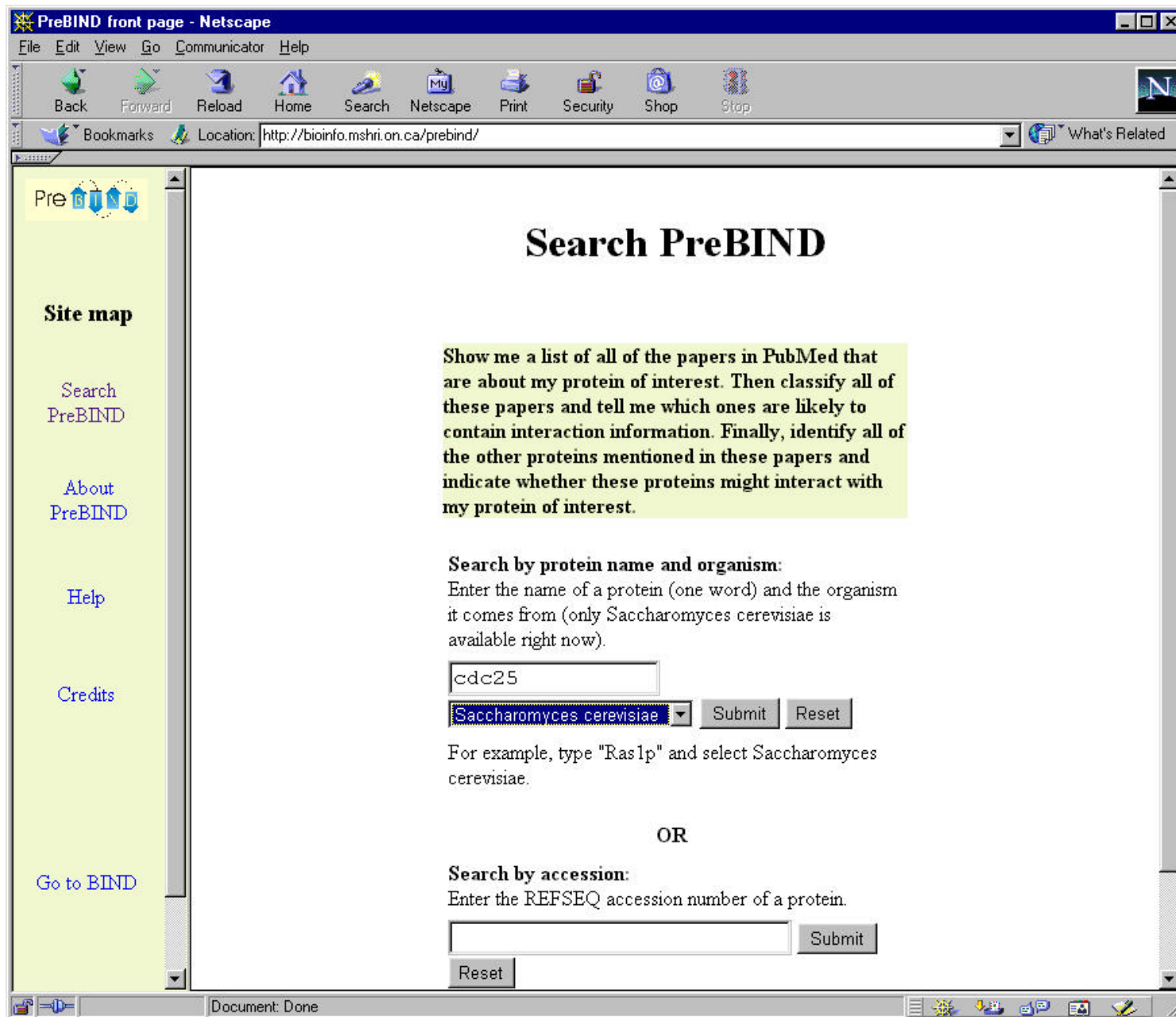
How many names in an abstract?

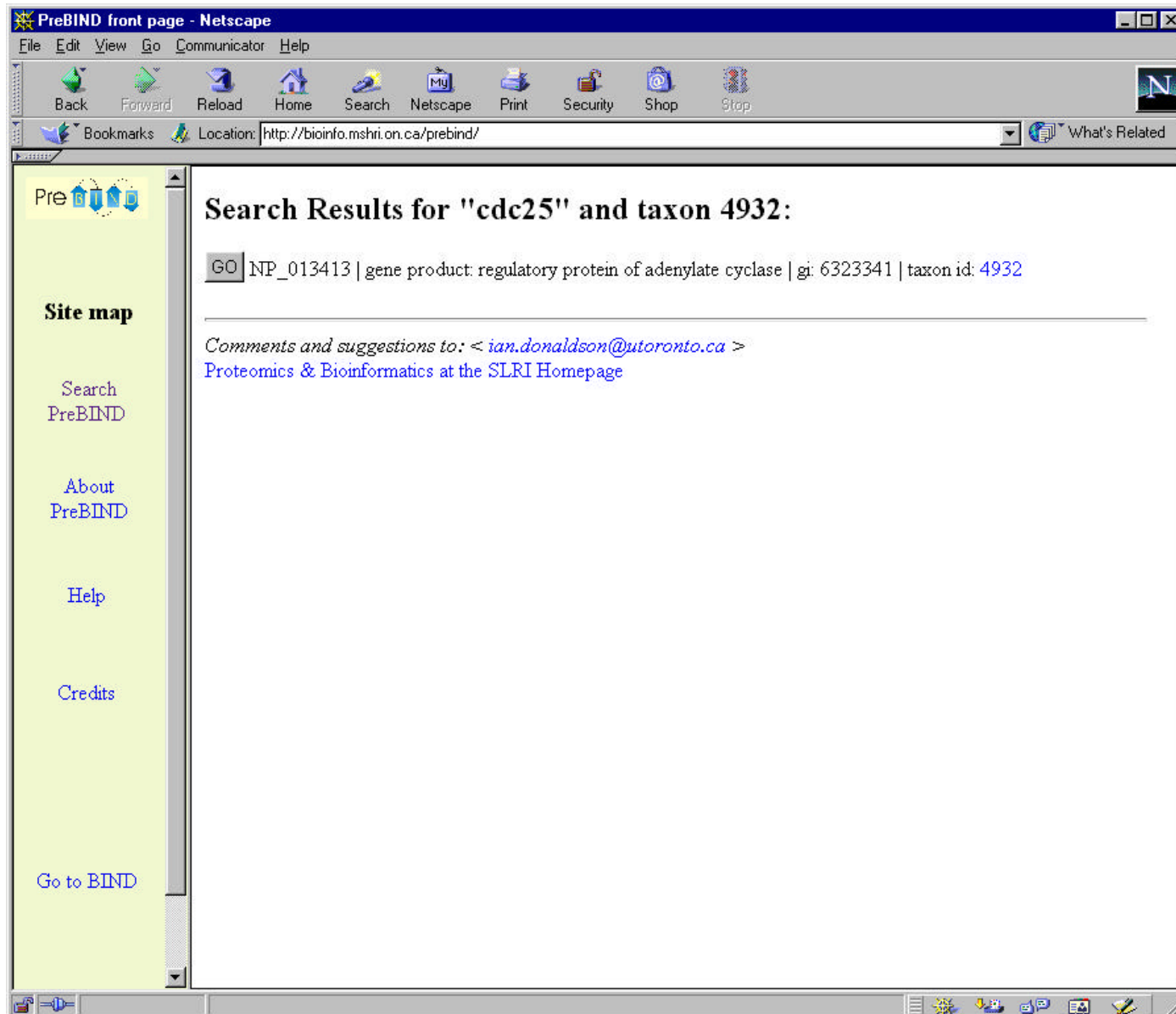


PreBIND numbers for yeast

Accession database unique proteins	6, 230
Search database (names)	11, 575
Total <i>S. cerevisiae</i> papers in Medline	42, 070
Results database (papers)	17, 043
Papers with 2 or more names	9, 361
>2 name interaction papers (svm)	2, 196
Potential interactions (score > 0)	1, 431








PreBIND front page - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://bioinfo.mshri.on.ca/prebind/> What's Related

Pre 

Site map

[Search PreBIND](#)

[About PreBIND](#)

[Help](#)

[Credits](#)

[Go to BIND](#)

Part 1 of 3: Protein description and name list

protein name: regulatory protein of adenylate cyclase

encoding locus name: CDC25

source database: REFSEQ

accession number: NP_013413

gi of protein: 6323341 ([SeqHound](#)) ([NCBI](#))

taxon id: [4932](#)

The following table lists the name(s) used to find the above gene product in the literature.

Search	Date last searched	Number of results found
CDC25	Jul 24, 2000	168
Cdc25p	Jul 24, 2000	18
Ctn1p	Jul 24, 2000	0
CTN1	Jul 24, 2000	0

Part 2 of 3: Interaction literature and intersecting names

The above names were used to search the literature for your protein of interest. Each section beginning with a blue triangle indicates a single PubMed abstract that was classified as containing interaction data. For each of these

Document: Done

PreBIND front page - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://bioinfo.mshri.on.ca/prebind/> What's Related

Part 2 of 3: Interaction literature and intersecting names

The above names were used to search the literature for your protein of interest. Each section beginning with a blue bar indicates a single PubMed abstract that was classified as containing interaction data. For each of these "interaction" papers, the other gene product names present in the title and abstract are listed. Natural language processing was used to determine if any of the named gene products was likely to be involved in an interaction with your protein of interest. Use the "GO" button to review all potential, pairwise interactions in the paper.

PubMed abstract [10724483](#) was returned by a search for "CDC25" and may contain interaction information (score = 0.100188).

Name(s) appearing in abstract	Description	Involved in an interaction?	Review this paper
HEM13	Coproporphyrinogen III oxidase	Unknown	GO
TIR1	Tir1p	Unknown	GO
GPA2/Gpa2p	nucleotide binding regulatory protein	Unknown	GO
HOG1	mitogen-activated protein kinase (MAP kinase)	Unknown	GO
GRE1	repressor activator protein	Unknown	GO

Pre

Site map

[Search PreBIND](#)

[About PreBIND](#)

[Help](#)

[Credits](#)

[Go to BIND](#)

PreBIND front page - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://bioinfo.mshri.on.ca/prebind/> What's Related

Site map

Search PreBIND

About PreBIND

Help

Credits

Go to BIND

Overview
Help | FAQ
New/Noteworthy

PubMed Services
Journal Browser
MeSH Browser
Single Citation
Matcher
Batch Citation
Matcher
Clinical Queries
Cubby

Related Resources
Order Documents
Grateful Med
Consumer Health

□ 1: *Mol Microbiol* 1999 Sep;33(5):904-18

online

Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast *Saccharomyces cerevisiae*.

Thevelein JM, de Winde JH

Laboratorium voor Moleculaire Celbiologie, Katholieke Universiteit Leuven, Kardinaal Mercierlaan 92, B-3001 Leuven-Heverlee, Flanders, Belgium. johan.thevelein@bio.kuleuven.ac.be

The cAMP-protein kinase A (PKA) pathway in the yeast *Saccharomyces cerevisiae* plays a major role in the control of metabolism, stress resistance and proliferation, in particular in connection with the available nutrient conditions. Extensive information has been obtained on the core section of the pathway, i.e. Cdc25, Ras, adenylate cyclase, PKA, and on components interacting directly with this core section, such as the Ira proteins, Cap/Srv2 and the two cAMP phosphodiesterases. Recent work has now started to reveal upstream regulatory components and downstream targets of

mentioned in the abstract (see the column "Involved in an interaction?" and the PreBIND Help page for more details). Use the "GO" button to submit feedback to PreBIND. Your input is appreciated.

Molecule A			Molecule B			Review	
Name(s) appearing in abstract	Description	Taxon	Name(s) appearing in abstract	Description	Taxon	Involved in an interaction?	Rev sub-inter
CLN3	G(sub)1 cyclin	4932	CDC28	protein kinase catalytic subunit	4932	Unknown Not reviewed.	GO

Related Articles, Books, LinkOut


Document: Done

PreBIND front page - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://bioinfo.mshri.on.ca/prebind/> What's Related



Site map

[Search PreBIND](#)

[About PreBIND](#)

[Help](#)

[Credits](#)

[Go to BIND](#)

Entrez PubMed
Overview
Help | FAQ
New/Noteworthy

PubMed Services
Journal Browser
MeSH Browser

☐ 1: *Mol Microbiol* 1999 Sep;33(5):904-18

online

Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast *Saccharomyces cerevisiae*.

[Related Articles, Books, LinkOut](#)

Review a potential interaction and make a submission to BIND

This form allows you to submit a simple Biomolecular Interaction Network Database (BIND) record. Information such as binding site location and experimental evidence cannot be entered using this form. If you cannot describe the interaction using the gene info ("gi") numbers listed below or if you want to describe an interaction involving a molecular complex, small molecule or RNA then you must enter this information directly at the [BIND site](#).

Step 1: Is there an interaction between CLN3 and CDC28 (either protein-protein or protein-DNA) described in the above PubMed paper 10476026?

☐ No, an interaction between these molecules is not described in this abstract. Skip to the last step below and press submit.

☐ Yes, an interaction between these molecules is described in this abstract. Continue with steps below and then press submit.

Step 2: Review and edit the description for molecule A.

Enter a short label for molecule A:

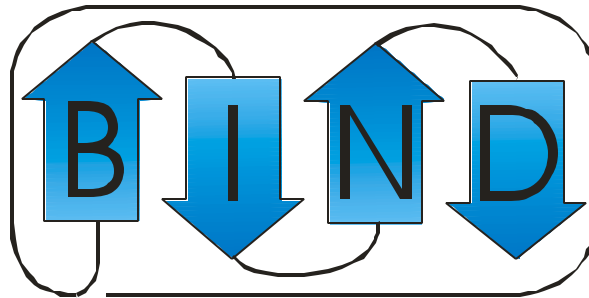
Choose the most appropriate name if more than one is presented

Description of Molecule A:

Document: Done

PubBIND - A public database of Interactions

- Biomolecular Interaction Network Database
- Public database, a submission site modeled after GenBank



binddb.org

How can I use BIND?

- **Public** www.bindddb.org (PubBIND)
- **Query** and explore the database.
- **Submit** data using the on-line forms and PreBIND.
- **Download** data, specifications and software.
- **Contribute** to the software base under open-source under the GNU Public License...

Without whom this would not be possible...

- **BIND & SeqHound**

- Gary Bader
- Ian Donaldson
- Cheryl Wolting
- Howard Feldman
- Katerina Michalikova
- Ruth Isserlin
- Berivan Baskin



- **BIND Collaborators**

- Tony Pawson, SLRI, UofT
- Viki Lay, MDS-Ocata
- Francis Ouellette, CMMT, UBC
- Christoph Sensen, NRC-IMB
- Joel Martin, NRC-IIT
- Jim Ostell, NCBI IEB
- NCBI - Structure Group
 - Steve Bryant, Tom Madej, Aron Marchler-Bauer, Lewis Geer, and more...

- **BIND - Licensed under the GNU Public license jointly by**

- MDS Proteomics
- Samuel Lunenfeld Research Inst.